

TRAITE DE COOPERATION EN MATIERE DE BREVETS

PCT

NOTIFICATION D'ELECTION

(règle 61.2 du PCT)

Expéditeur: le BUREAU INTERNATIONAL

Destinataire:

Assistant Commissioner for Patents
United States Patent and Trademark
Office
Box PCT
Washington, D.C.20231
ETATS-UNIS D'AMERIQUE

en sa qualité d'office élu

Date d'expédition (jour/mois/année) 24 juillet 2000 (24.07.00)	
Demande internationale no PCT/FR99/02941	Référence du dossier du déposant ou du mandataire CMG598/25
Date du dépôt international (jour/mois/année) 26 novembre 1999 (26.11.99)	Date de priorité (jour/mois/année) 26 novembre 1998 (26.11.98)
Déposant BEAUVILLAIN, Jean-Claude etc	

1. L'office désigné est avisé de son élection qui a été faite:

☒ dans la demande d'examen préliminaire international présentée à l'administration chargée de l'examen préliminaire international le:

16 juin 2000 (16.06.00)

☐ dans une déclaration visant une élection ultérieure déposée auprès du Bureau international le:

2. L'élection ☒ a été faite

☐ n'a pas été faite

avant l'expiration d'un délai de 19 mois à compter de la date de priorité ou, lorsque la règle 32 s'applique, dans le délai visé à la règle 32.2b).

Bureau international de l'OMPI 34, chemin des Colombettes 1211 Genève 20, Suisse	Fonctionnaire autorisé <p style="text-align: center;">Antonia Muller</p>
no de télécopieur: (41-22) 740.14.35	no de téléphone: (41-22) 338.83.38

CORRECTED
VERSION*

PCT

WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau



INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 6 : C12N 15/31, C07K 14/27, A01H 1/00, 3/00		A3	(11) International Publication Number: WO 98/54214 (43) International Publication Date: 3 December 1998 (03.12.98)
(21) International Application Number: PCT/US98/10874 (22) International Filing Date: 28 May 1998 (28.05.98) (30) Priority Data: 60/048,109 30 May 1997 (30.05.97) US (71) Applicants: CORNELL RESEARCH FOUNDATION, INC. [US/US]; Suite 105, 20 Thornwood Drive, Ithaca, NY 14850 (US). EDEN BIOSCIENCE CORPORATION [US/US]; 11816 North Creek Parkway N., Bothell, WA 98011-8205 (US). (72) Inventors: LABY, Ronald, J.; 1411 Haddon Street, Houston, TX 77006 (US). WEI, Zhong-Min; 8230 125th Court, Kirkland, WA 98034 (US). BEER, Steven, V.; 211 Hudson Street, Ithaca, NY 14850 (US). (74) Agents: GOLDMAN, Michael, L. et al.; Nixon, Hargrave, Devans & Doyle LLP, Clinton Square, P.O. Box 1051, Rochester, NY 14603 (US).		(81) Designated States: AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE, DK, EE, ES, FI, GB, GE, GH, GM, GW, HU, ID, IL, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, UA, UG, UZ, VN, YU, ZW, ARIPO patent (GH, GM, KE, LS, MW, SD, SZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, ML, MR, NE, SN, TD, TG). Published <i>With international search report.</i> (88) Date of publication of the international search report: 4 March 1999 (04.03.99)	
(54) Title: HYPERSENSITIVE RESPONSE ELICITOR FRAGMENTS AND USES THEREOF			
(57) Abstract <p>The present invention is directed to isolated fragments of an <i>Erwinia</i> hypersensitive response elicitor protein or polypeptide which fragments elicit a hypersensitive response in plants. Also disclosed are isolated DNA molecules which encode the <i>Erwinia</i> hypersensitive response eliciting fragment. Isolated fragments of hypersensitive response elicitor proteins or polypeptides, which elicit a hypersensitive response, and the isolated DNA molecules that encode them can be used to impart disease resistance to plants, to enhance plant growth, and/or to control insects on plants. This can be achieved by applying the hypersensitive response eliciting fragments in a non-infectious form to plants or plant seeds under conditions effective to impart disease resistance, to enhance plant growth, and/or to control insects on plants or plants grown from the plant seeds. Alternatively, transgenic plants or plant seeds transformed with a DNA molecule encoding a hypersensitive response eliciting fragment can be provided and the transgenic plants or plants resulting from the transgenic plant seeds are grown under conditions effective to impart disease resistance, to enhance plant growth, and/or to control insects on plants or plants grown from the plant seeds.</p>			

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	ES	Spain	LS	Lesotho	SI	Slovenia
AM	Armenia	FI	Finland	LT	Lithuania	SK	Slovakia
AT	Austria	FR	France	LU	Luxembourg	SN	Senegal
AU	Australia	GA	Gabon	LV	Latvia	SZ	Swaziland
AZ	Azerbaijan	GB	United Kingdom	MC	Monaco	TD	Chad
BA	Bosnia and Herzegovina	GE	Georgia	MD	Republic of Moldova	TG	Togo
BB	Barbados	GH	Ghana	MG	Madagascar	TJ	Tajikistan
BE	Belgium	GN	Guinea	MK	The former Yugoslav Republic of Macedonia	TM	Turkmenistan
BF	Burkina Faso	GR	Greece			TR	Turkey
BG	Bulgaria	HU	Hungary	ML	Mali	TT	Trinidad and Tobago
BJ	Benin	IE	Ireland	MN	Mongolia	UA	Ukraine
BR	Brazil	IL	Israel	MR	Mauritania	UG	Uganda
BY	Belarus	IS	Iceland	MW	Malawi	US	United States of America
CA	Canada	IT	Italy	MX	Mexico	UZ	Uzbekistan
CF	Central African Republic	JP	Japan	NE	Niger	VN	Viet Nam
CG	Congo	KE	Kenya	NL	Netherlands	YU	Yugoslavia
CH	Switzerland	KG	Kyrgyzstan	NO	Norway	ZW	Zimbabwe
CI	Côte d'Ivoire	KP	Democratic People's Republic of Korea	NZ	New Zealand		
CM	Cameroon	KR	Republic of Korea	PL	Poland		
CN	China			PT	Portugal		
CU	Cuba	KZ	Kazakhstan	RO	Romania		
CZ	Czech Republic	LC	Saint Lucia	RU	Russian Federation		
DE	Germany	LI	Liechtenstein	SD	Sudan		
DK	Denmark	LK	Sri Lanka	SE	Sweden		
EE	Estonia	LR	Liberia	SG	Singapore		

HYPERSENSITIVE RESPONSE ELICITOR FRAGMENTS AND USES THEREOF

This application claims benefit of U.S. Patent Application Serial
5 No. 60/048,109.

FIELD OF THE INVENTION

The present invention relates to fragments of a hypersensitive response elicitor
10 which fragments elicit a hypersensitive response and uses thereof.

BACKGROUND OF THE INVENTION

Interactions between bacterial pathogens and their plant hosts generally fall
15 into two categories: (1) compatible (pathogen-host), leading to intercellular bacterial growth, symptom development, and disease development in the host plant; and (2) incompatible (pathogen-nonhost), resulting in the hypersensitive response, a particular type of incompatible interaction occurring, without progressive disease symptoms. During compatible interactions on host plants, bacterial populations increase
20 dramatically and progressive symptoms occur. During incompatible interactions, bacterial populations do not increase, and progressive symptoms do not occur.

The hypersensitive response is a rapid, localized necrosis that is associated with the active defense of plants against many pathogens (Kiraly, Z., "Defenses Triggered by the Invader: Hypersensitivity," pages 201-224 in: Plant Disease: An
25 Advanced Treatise, Vol. 5, J.G. Horsfall and E.B. Cowling, ed. Academic Press New York (1980); Klement, Z., "Hypersensitivity," pages 149-177 in: Phytopathogenic Prokaryotes, Vol. 2, M.S. Mount and G.H. Lacy, ed. Academic Press, New York (1982)). The hypersensitive response elicited by bacteria is readily observed as a tissue collapse if high concentrations ($\geq 10^7$ cells/ml) of a limited host-range pathogen
30 like *Pseudomonas syringae* or *Erwinia amylovora* are infiltrated into the leaves of nonhost plants (necrosis occurs only in isolated plant cells at lower levels of inoculum) (Klement, Z., "Rapid Detection of Pathogenicity of Phytopathogenic Pseudomonads," Nature 199:299-300; Klement, et al., "Hypersensitive Reaction Induced by Phytopathogenic Bacteria in the Tobacco Leaf," Phytopathology 54:474-

- 2 -

- 477 (1963); Turner, et al., "The Quantitative Relation Between Plant and Bacterial Cells Involved in the Hypersensitive Reaction," Phytopathology 64:885-890 (1974); Klement, Z., "Hypersensitivity," pages 149-177 in Phytopathogenic Prokaryotes, Vol. 2., M.S. Mount and G.H. Lacy, ed. Academic Press, New York (1982)). The capacities to elicit the hypersensitive response in a nonhost and be pathogenic in a host appear linked. As noted by Klement, Z., "Hypersensitivity," pages 149-177 in Phytopathogenic Prokaryotes, Vol. 2., M.S. Mount and G.H. Lacy, ed. Academic Press, New York, these pathogens also cause physiologically similar, albeit delayed, necroses in their interactions with compatible hosts. Furthermore, the ability to produce the hypersensitive response or pathogenesis is dependent on a common set of genes, denoted *hrp* (Lindgren, P.B., et al., "Gene Cluster of *Pseudomonas syringae* pv. 'phaseolicola' Controls Pathogenicity of Bean Plants and Hypersensitivity on Nonhost Plants," J. Bacteriol. 168:512-22 (1986); Willis, D.K., et al., "*hrp* Genes of Phytopathogenic Bacteria," Mol. Plant-Microbe Interact. 4:132-138 (1991)). Consequently, the hypersensitive response may hold clues to both the nature of plant defense and the basis for bacterial pathogenicity.

- The *hrp* genes are widespread in gram-negative plant pathogens, where they are clustered, conserved, and in some cases interchangeable (Willis, D.K., et al., "*hrp* Genes of Phytopathogenic Bacteria," Mol. Plant-Microbe Interact. 4:132-138 (1991); Bonas, U., "*hrp* Genes of Phytopathogenic Bacteria," pages 79-98 in: Current Topics in Microbiology and Immunology: Bacterial Pathogenesis of Plants and Animals - Molecular and Cellular Mechanisms, J.L. Dangel, ed. Springer-Verlag, Berlin (1994)). Several *hrp* genes encode components of a protein secretion pathway similar to one used by *Yersinia*, *Shigella*, and *Salmonella* spp. to secrete proteins essential in animal diseases (Van Gijsegem, et al., "Evolutionary Conservation of Pathogenicity Determinants Among Plant and Animal Pathogenic Bacteria," Trends Microbiol. 1:175-180 (1993)). In *E. amylovora*, *P. syringae*, and *P. solanacearum*, *hrp* genes have been shown to control the production and secretion of glycine-rich, protein elicitors of the hypersensitive response (He, S.Y., et al. "*Pseudomonas Syringae* pv. *Syringae* HarpinPss: a Protein that is Secreted via the Hrp Pathway and Elicits the Hypersensitive Response in Plants," Cell 73:1255-1266 (1993), Wei, Z.-H., et al., "*HrpI* of *Erwinia amylovora* Functions in Secretion of Harpin and is a Member of a

- 3 -

New Protein Family," J. Bacteriol. 175:7958-7967 (1993); Arlat, M. et al. "PopA1, a Protein Which Induces a Hypersensitive-like Response on Specific Petunia Genotypes, is Secreted via the Hrp Pathway of *Pseudomonas solanacearum*," EMBO J. 13:543-553 (1994)).

5 The first of these proteins was discovered in *E. amylovora* Ea321, a bacterium that causes fire blight of rosaceous plants, and was designated harpin (Wei, Z.-M., et al, "Harpin, Elicitor of the Hypersensitive Response Produced by the Plant Pathogen *Erwinia amylovora*," Science 257:85-88 (1992)). Mutations in the encoding *hrpN* gene revealed that harpin is required for *E. amylovora* to elicit a hypersensitive
10 response in nonhost tobacco leaves and incite disease symptoms in highly susceptible pear fruit. The *P. solanacearum* GMI1000 PopA1 protein has similar physical properties and also elicits the hypersensitive response in leaves of tobacco, which is not a host of that strain (Arlat, et al. "PopA1, a Protein Which Induces a Hypersensitive-like Response on Specific Petunia Genotypes, is Secreted via the Hrp
15 Pathway of *Pseudomonas solanacearum*," EMBO J. 13:543-53 (1994)). However, *P. solanacearum* *popA* mutants still elicit the hypersensitive response in tobacco and incite disease in tomato. Thus, the role of these glycine-rich hypersensitive response elicitors can vary widely among gram-negative plant pathogens.

Other plant pathogenic hypersensitive response elicitors have been isolated,
20 cloned, and sequenced. These include: *Erwinia chrysanthemi* (Bauer, et. al., "Erwinia chrysanthemi Harpin_{Ech}: Soft-Rot Pathogenesis," MPMI 8(4): 484-91 (1995)); *Erwinia carotovora* (Cui, et. al., "The RsmA⁻ Mutants of *Erwinia carotovora* subsp. *carotovora* Strain Ecc71 Overexpress *hrpN*_{Ecc} and Elicit a Hypersensitive Reaction-like Response in Tobacco Leaves," MPMI 9(7): 565-73 (1996)); *Erwinia*
25 *stewartii* (Ahmad, et. al., "Harpin is not Necessary for the Pathogenicity of *Erwinia stewartii* on Maize," 8th Int'l. Cong. Molec. Plant-Microb. Inter. July 14-19, 1996 and Ahmad, et. al., "Harpin is not Necessary for the Pathogenicity of *Erwinia stewartii* on Maize," Ann. Mtg. Am. Phytopath. Soc. July 27-31, 1996); and *Pseudomonas syringae* pv. *syringae* (WO 94/26782 to Cornell Research Foundation, Inc.).

30 The present invention seeks to identify fragments of hypersensitive response elicitor proteins or polypeptides, which fragments elicit a hypersensitive response, and uses of such fragments.

SUMMARY OF THE INVENTION

The present invention is directed to an isolated fragment of an *Erwinia*
5 hypersensitive response elicitor protein or polypeptide where the fragment elicits a
hypersensitive response in plants. Also disclosed are isolated DNA molecules which
encode such fragments.

The fragments of hypersensitive response elicitors can be used to impart
disease resistance to plants, to enhance plant growth, and/or to control insects. This
10 involves applying the fragments in a non-infectious form to plants or plant seeds
under conditions effective to impart disease resistance, to enhance plant growth,
and/or to control insects on plants or plants grown from the plant seeds.

As an alternative to applying the fragments to plants or plant seeds in order to
impart disease resistance, to enhance plant growth, and/or to control insects on plants,
15 transgenic plants or plant seeds can be utilized. When utilizing transgenic plants, this
involves providing a transgenic plant transformed with a DNA molecule encoding a
fragment of a hypersensitive response elicitor protein or polypeptide which fragments
elicit a hypersensitive response in plants and growing the plant under conditions
effective to impart disease resistance, to enhance plant growth, and/or to control
20 insects in the plants or plants grown from the plant seeds. Alternatively, a transgenic
plant seed transformed with the DNA molecule encoding such a fragment can be
provided and planted in soil. A plant is then propagated under conditions effective to
impart disease resistance, to enhance plant growth, and/or to control insects on plants
or plants grown from the plant seeds.

25

BRIEF DESCRIPTION OF THE DRAWINGS

Figure 1 shows a deletion and proteolysis analysis for the *Erwinia amylovora*
hypersensitive response elicitor (i.e. harpin). A is the name of the harpin fragment. B
30 is the length of the fragment in amino acid residues. C indicates whether detectable
protein is produced. D states whether there is hypersensitive response (i.e., HR)
eliciting activity. The solid line indicates that there are additional amino acids which
are not harpin encoded, while the dashed line indicates the portion of the harpin that is

- 5 -

deleted. The numbers above the fragments in the box represent the amino acid residue present at the end of a given fragment; residue #1 is the N-terminus, and residue #403 is the C-terminus.

Figure 2 is a Western blot illustrating specific secretion of harpin_{Ea}, but not harpin_{Ea}C31. Lane A, Ea273(pGP1-2) CFEP; lane B, Ea273(pGP1-2)(pCPP1104) CFEP; lane C, *E. coli* DH5α (pCPP1107) CFEP harpin size standard; lane D, BioRad low range molecular weight markers; lane E, Ea273(pGP1-2) supernatant; lane F, Ea273(pGP1 2)(pCPP1104) supernatant. The blot was probed with an anti-harpin_{Ea} polyclonal antibody.

Figure 3 is an HR assay on tobacco leaf infiltrated as follows: (1) A, harpin_{Ea} + raspberry IF; (2) B, harpin_{Ea} + apple IF; (3) C, harpin_{Ea} + tobacco IF; (4) D, harpin_{Ea} + endoproteinase Glu-C; (5) E, harpin_{Ea} + trypsin; (6) F, harpin_{Ea}; (7) G, tobacco IF; (8) H, endoproteinase Glu-C; (9) I, trypsin; and (10) J, harpin_{Ea}. IF refers to intracellular fluids.

Figure 4 shows the digestion of harpin with endoproteinase Glu-C. Lane A is harpin; Lane B is harpin + endoproteinase Glu-C; Lane C is BioRad low range molecular weight markers.

Figure 5A shows the proteolysis of harpin. Coomassie blue stained polyacrylamide gel was loaded as follows: A, BioRad low range molecular weight markers; B, IF-apple; C, IF-raspberry; D, IF-tobacco; E, harpin_{Ea}; F, harpin_{Ea} + IF-apple; G, harpin_{Ea} + IF-raspberry; H, harpin_{Ea} + IF-tobacco.

Figure 5B shows a Coomassie Blue stained polyacrylamide gel loaded as follows: A, IF-tobacco; B, IF-tobacco + harpin_{Ea}; C, harpin_{Ea}; D, BioRad low range molecular weight markers; E, IF-tobacco + harpin_{Ea} + PMSF. HR-eliciting activity of the sample following proteolysis is denoted below the gel.

Figure 5C depicts whether proteolytic activity is present in IF from all plants tested. Intercellular fluid harvested from several plants was analyzed by PAGE in a gel containing 0.1% copolymerized gelatin. After washing to remove SDS and incubation to allow proteolysis of gelatin, the gels were stained to demonstrate the presence of gelatinolytic activity. A, IF-apple; B, IF-tobacco; C, IF-cotoneaster; D, BioRad mw; E, endoproteinase Glu-C; and F, ground leaf extract-tobacco.

- 6 -

Figure 6 shows the refractionation of elicitor-active peptides following proteolysis of harpin_{Ea} by tobacco IF. Absorbance was measured at 210 nm. Peak 1 contains peptides P91 and P95; peak 2 contains peptides P65 and P69.

Figure 7 shows the predicted proteolytic cleavage sites within harpin of several tested proteinases, and the effect of these cleavages on activity of active harpin fragments. Residues potentially important for HR-eliciting activity, based on the loss of activity following further cleavage, are indicated by upward-pointing arrows at bottom.

Figure 8 shows the similarities near N-termini among harpins of *Erwinia* spp. Underlined residues are present (identical or similar) in at least four out of the five proteins examined. Nine out of the first 26 residues are conserved in this manner.

Figures 9 A-B show Kyte-Doolittle hydropathy plots of bacterial HR-eliciting proteins. Ea, *E. amylovora* EA321; Est, *E. stewartii* DC283; Ech, *E. chrysanthemi* AC4150; Ecc, *E. carotovora* subsp. *carotovora*; Rs, *R. solanacearum*; Pss, *P. syringae* pv. *syringae*.

Figure 10 shows truncated proteins of the hypersensitive response elicitor protein or polypeptide.

Figure 11 shows a list of synthesized oligonucleotide primers for construction of truncated harpin proteins. N represents the N-terminus (5' region), and C represents the C-terminus (3' region). The primers correspond to the indicated sequence identification numbers for the present application: N1 (SEQ. ID. No. 1), N76 (SEQ. ID. No. 2), N99 (SEQ. ID. No. 3), N105 (SEQ. ID. No. 4), N110 (SEQ. ID. No. 5), N137 (SEQ. ID. No. 6), N150 (SEQ. ID. No. 7), N169 (SEQ. ID. No. 8), N210 (SEQ. ID. No. 9), N267 (SEQ. ID. No. 10), N343 (SEQ. ID. No. 11), C75 (SEQ. ID. No. 12), C104 (SEQ. ID. No. 13), C168 (SEQ. ID. No. 14), C180 (SEQ. ID. No. 15), C204 (SEQ. ID. No. 16), C209 (SEQ. ID. No. 17), C266 (SEQ. ID. No. 18), C342 (SEQ. ID. No. 19), and C403 (SEQ. ID. No. 20).

DETAILED DESCRIPTION OF THE INVENTION

The present invention is directed to isolated fragments of a hypersensitive response elicitor protein or polypeptide where the fragments elicit a hypersensitive response in plants. Also disclosed are DNA molecules encoding such fragments as

- 7 -

well as expression systems, host cells, and plants containing such molecules. Uses of the fragments themselves and the DNA molecules encoding them are disclosed.

The fragments of hypersensitive response elicitor polypeptides or proteins according to the present invention are derived from hypersensitive response elicitor polypeptides or proteins of a wide variety of fungal and bacterial pathogens. Such polypeptides or proteins are able to elicit local necrosis in plant tissue contacted by the elicitor. Examples of suitable bacterial sources of polypeptide or protein elicitors include *Erwinia*, *Pseudomonas*, and *Xanthomonas* species (e.g., the following bacteria: *Erwinia amylovora*, *Erwinia chrysanthemi*, *Erwinia stewartii*, *Erwinia carotovora*, *Pseudomonas syringae*, *Pseudomonas solanacearum*, *Xanthomonas campestris*, and mixtures thereof).

An example of a fungal source of a hypersensitive response elicitor protein or polypeptide is *Phytophthora*. Suitable species of *Phytophthora* include *Phytophthora parasitica*, *Phytophthora cryptogea*, *Phytophthora cinnamomi*, *Phytophthora capsici*, *Phytophthora megasperma*, and *Phytophthora citrophthora*.

The hypersensitive response elicitor polypeptide or protein from *Erwinia chrysanthemi* has an amino acid sequence corresponding to SEQ. ID. No. 21 as follows:

20	Met Gln Ile Thr Ile Lys Ala His Ile Gly Gly Asp Leu Gly Val Ser	1	5	10	15
	Gly Leu Gly Ala Gln Gly Leu Lys Gly Leu Asn Ser Ala Ala Ser Ser	20	25	30	
25	Leu Gly Ser Ser Val Asp Lys Leu Ser Ser Thr Ile Asp Lys Leu Thr	35	40	45	
	Ser Ala Leu Thr Ser Met Met Phe Gly Gly Ala Leu Ala Gln Gly Leu	50	55	60	
	Gly Ala Ser Ser Lys Gly Leu Gly Met Ser Asn Gln Leu Gly Gln Ser	65	70	75	80
30	Phe Gly Asn Gly Ala Gln Gly Ala Ser Asn Leu Leu Ser Val Pro Lys	85	90	95	
	Ser Gly Gly Asp Ala Leu Ser Lys Met Phe Asp Lys Ala Leu Asp Asp	100	105	110	
35	Leu Leu Gly His Asp Thr Val Thr Lys Leu Thr Asn Gln Ser Asn Gln	115	120	125	

- 8 -

	Leu	Ala	Asn	Ser	Met	Leu	Asn	Ala	Ser	Gln	Met	Thr	Gln	Gly	Asn	Met	
	130						135						140				
	Asn	Ala	Phe	Gly	Ser	Gly	Val	Asn	Asn	Ala	Leu	Ser	Ser	Ile	Leu	Gly	
	145					150					155					160	
5	Asn	Gly	Leu	Gly	Gln	Ser	Met	Ser	Gly	Phe	Ser	Gln	Pro	Ser	Leu	Gly	
					165					170					175		
	Ala	Gly	Gly	Leu	Gln	Gly	Leu	Ser	Gly	Ala	Gly	Ala	Phe	Asn	Gln	Leu	
				180					185					190			
10	Gly	Asn	Ala	Ile	Gly	Met	Gly	Val	Gly	Gln	Asn	Ala	Ala	Leu	Ser	Ala	
		195					200						205				
	Leu	Ser	Asn	Val	Ser	Thr	His	Val	Asp	Gly	Asn	Asn	Arg	His	Phe	Val	
	210						215					220					
	Asp	Lys	Glu	Asp	Arg	Gly	Met	Ala	Lys	Glu	Ile	Gly	Gln	Phe	Met	Asp	
	225					230					235					240	
15	Gln	Tyr	Pro	Glu	Ile	Phe	Gly	Lys	Pro	Glu	Tyr	Gln	Lys	Asp	Gly	Trp	
					245					250					255		
	Ser	Ser	Pro	Lys	Thr	Asp	Asp	Lys	Ser	Trp	Ala	Lys	Ala	Leu	Ser	Lys	
				260					265					270			
20	Pro	Asp	Asp	Asp	Gly	Met	Thr	Gly	Ala	Ser	Met	Asp	Lys	Phe	Arg	Gln	
		275						280					285				
	Ala	Met	Gly	Met	Ile	Lys	Ser	Ala	Val	Ala	Gly	Asp	Thr	Gly	Asn	Thr	
		290						295				300					
	Asn	Leu	Asn	Leu	Arg	Gly	Ala	Gly	Gly	Ala	Ser	Leu	Gly	Ile	Asp	Ala	
	305					310					315					320	
25	Ala	Val	Val	Gly	Asp	Lys	Ile	Ala	Asn	Met	Ser	Leu	Gly	Lys	Leu	Ala	
					325					330					335		
	Asn	Ala															

30 This hypersensitive response elicitor polypeptide or protein has a molecular weight of 34 kDa, is heat stable, has a glycine content of greater than 16%, and contains substantially no cysteine. The *Erwinia chrysanthemi* hypersensitive response elicitor polypeptide or protein is encoded by a DNA molecule having a nucleotide sequence corresponding to SEQ. ID. No. 22 as follows:

35

CGATTTTACC	CGGGTGAACG	TGCTATGACC	GACAGCATCA	CGGTATTCCA	CACCGTTACG	60
GCGTTTATGG	CCGCGATGAA	CCGGCATCAG	GCGGCGCGCT	GGTCGCCGCA	ATCCGGCGTC	120

- 9 -

	GATCTGGTAT TTCAGTTTGG GGACACCGGG CGTGAACCTCA TGATGCAGAT TCAGCCGGGG	180
	CAGCAATATC CCGGCATGTT GCGCACGCTG CTCGCTCGTC GTTATCAGCA GGCGGCAGAG	240
	TGCGATGGCT GCCATCTGTG CCTGAACGGC AGCGATGTAT TGATCCTCTG GTGGCCGCTG	300
	CCGTCGGATC CCGGCAGTTA TCCGCAGGTG ATCGAACGTT TGTTTGAACT GGCGGGAATG	360
5	ACGTTGCCGT CGCTATCCAT AGCACCGACG GCGCGTCCGC AGACAGGGAA CGGACGCGCC	420
	CGATCATTAA GATAAAGGCG GCTTTTTTTA TTGCAAAACG GTAACGGTGA GGAACCGTTT	480
	CACCGTCGGC GTCACCTCAGT AACAAGTATC CATCATGATG CCTACATCGG GATCGGCGTG	540
	GGCATCCGTT GCAGATACTT TTGCGAACAC CTGACATGAA TGAGGAAACG AAATTATGCA	600
	AATTACGATC AAAGCGCACA TCGGCGGTGA TTTGGGCGTC TCCGGTCTGG GGCTGGGTGC	660
10	TCAGGGACTG AAAGGACTGA ATTCGCGGGC TTCATCGCTG GGTTCACGG TGGATAAACT	720
	GAGCAGCACC ATCGATAAGT TGACCTCCGC GCTGACTTCG ATGATGTTTG GCGGCGCGCT	780
	GGCGCAGGGG CTGGGCGCCA GCTCGAAGGG GCTGGGGATG AGCAATCAAC TGGGCCAGTC	840
	TTTCGGCAAT GGC GCGCAGG GTGCGAGCAA CCTGCTATCC GTACCGAAAT CCGGCGGCGA	900
	TGCGTTGTCA AAAATGTTTG ATAAAGCGCT GGACGATCTG CTGGGTCATG ACACCGTGAC	960
15	CAAGCTGACT AACCAGAGCA ACCAACTGGC TAATTCAATG CTGAACGCCA GCCAGATGAC	1020
	CCAGGGTAAT ATGAATGCGT TCGGCGCGG TGTGAACAAC GCACTGTCGT CCATTCTCGG	1080
	CAACGGTCTC GGCCAGTCGA TGAGTGGCTT CTCTCAGCCT TCTCTGGGGG CAGGCGGCTT	1140
	GCAGGGCCTG AGCGGCGCGG GTGCATTCAA CCAGTTGGGT AATGCCATCG GCATGGGCGT	1200
	GGGGCAGAAT GCTGCGCTGA GTGCGTTGAG TAACGTCAGC ACCCACGTAG ACGGTAACAA	1260
20	CCGCCACTTT GTAGATAAAG AAGATCGCGG CATGGCGAAA GAGATCGGCC AGTTTATGGA	1320
	TCAGTATCCG GAAATATTCG GTAAACCGGA ATACCAGAAA GATGGCTGGA GTTCGCCGAA	1380
	GACGGACGAC AAATCCTGGG CTAAAGCGCT GAGTAAACCG GATGATGACG GTATGACCGG	1440
	CGCCAGCATG GACAAATTCC GTCAGGCGAT GGGTATGATC AAAAGCGCGG TGGCGGGTGA	1500
	TACCGGCAAT ACCAACCTGA ACCTGCGTGG CGCGGGCGGT GCATCGCTGG GTATCGATGC	1560
25	GGCTGTCGTC GGCGATAAAA TAGCCAACAT GTCGCTGGGT AAGCTGGCCA ACGCCTGATA	1620
	ATCTGTGCTG GCCTGATAAA GCGGAAACGA AAAAAGAGAC GGGGAAGCCT GTCTCTTTTC	1680
	TTATTATGCG GTTTATGCGG TTACCTGGAC CGGTTAATCA TCGTCATCGA TCTGGTACAA	1740
	ACGCACATTT TCCCGTTCAT TCGCGTCGTT ACGCGCCACA ATCGCGATGG CATCTTCCTC	1800
	GTCGCTCAGA TTGCGCGGCT GATGGGGAAC GCCGGGTGGA ATATAGAGAA ACTCGCCGGC	1860

- 10 -

CAGATGGAGA CACGTCTGCG ATAAATCTGT GCCGTAACGT GTTCTATCC GCCCCTTTAG 1920
 CAGATAGATT GCGGTTTCGT AATCAACATG GTAATGCGGT TCCGCCTGTG CGCCGGCCGG 1980
 GATCACCACA ATATTCATAG AAAGCTGTCT TGCACCTACC GTATCGCGGG AGATACCGAC 2040
 AAAATAGGGC AGTTTTTGCG TGGTATCCGT GGGGTGTTCC GGCCTGACAA TCTTGAGTTG 2100
 5 GTTCGTCATC ATCTTTCTCC ATCTGGGCGA CCTGATCGGT T 2141

The hypersensitive response elicitor polypeptide or protein derived from
Erwinia amylovora has an amino acid sequence corresponding to SEQ. ID. No. 23 as
 10 follows:

Met Ser Leu Asn Thr Ser Gly Leu Gly Ala Ser Thr Met Gln Ile Ser
 1 5 10 15
 Ile Gly Gly Ala Gly Gly Asn Asn Gly Leu Leu Gly Thr Ser Arg Gln
 15 20 25 30
 Asn Ala Gly Leu Gly Gly Asn Ser Ala Leu Gly Leu Gly Gly Gly Asn
 35 40 45
 Gln Asn Asp Thr Val Asn Gln Leu Ala Gly Leu Leu Thr Gly Met Met
 50 55 60
 20 Met Met Met Ser Met Met Gly Gly Gly Gly Leu Met Gly Gly Gly Leu
 65 70 75 80
 Gly Gly Gly Leu Gly Asn Gly Leu Gly Gly Ser Gly Gly Leu Gly Glu
 85 90 95
 Gly Leu Ser Asn Ala Leu Asn Asp Met Leu Gly Gly Ser Leu Asn Thr
 100 105 110
 25 Leu Gly Ser Lys Gly Gly Asn Asn Thr Thr Ser Thr Thr Asn Ser Pro
 115 120 125
 Leu Asp Gln Ala Leu Gly Ile Asn Ser Thr Ser Gln Asn Asp Asp Ser
 130 135 140
 30 Thr Ser Gly Thr Asp Ser Thr Ser Asp Ser Ser Asp Pro Met Gln Gln
 145 150 155 160
 Leu Leu Lys Met Phe Ser Glu Ile Met Gln Ser Leu Phe Gly Asp Gly
 165 170 175
 Gln Asp Gly Thr Gln Gly Ser Ser Ser Gly Gly Lys Gln Pro Thr Glu
 180 185 190
 35 Gly Glu Gln Asn Ala Tyr Lys Lys Gly Val Thr Asp Ala Leu Ser Gly
 195 200 205

- 11 -

	Leu	Met	Gly	Asn	Gly	Leu	Ser	Gln	Leu	Leu	Gly	Asn	Gly	Gly	Leu	Gly	
	210						215					220					
	Gly	Gly	Gln	Gly	Gly	Asn	Ala	Gly	Thr	Gly	Leu	Asp	Gly	Ser	Ser	Leu	
	225					230					235					240	
5	Gly	Gly	Lys	Gly	Leu	Gln	Asn	Leu	Ser	Gly	Pro	Val	Asp	Tyr	Gln	Gln	
					245					250					255		
	Leu	Gly	Asn	Ala	Val	Gly	Thr	Gly	Ile	Gly	Met	Lys	Ala	Gly	Ile	Gln	
				260					265						270		
10	Ala	Leu	Asn	Asp	Ile	Gly	Thr	His	Arg	His	Ser	Ser	Thr	Arg	Ser	Phe	
			275					280					285				
	Val	Asn	Lys	Gly	Asp	Arg	Ala	Met	Ala	Lys	Glu	Ile	Gly	Gln	Phe	Met	
	290						295					300					
	Asp	Gln	Tyr	Pro	Glu	Val	Phe	Gly	Lys	Pro	Gln	Tyr	Gln	Lys	Gly	Pro	
	305					310					315					320	
15	Gly	Gln	Glu	Val	Lys	Thr	Asp	Asp	Lys	Ser	Trp	Ala	Lys	Ala	Leu	Ser	
					325					330					335		
	Lys	Pro	Asp	Asp	Asp	Gly	Met	Thr	Pro	Ala	Ser	Met	Glu	Gln	Phe	Asn	
				340					345					350			
20	Lys	Ala	Lys	Gly	Met	Ile	Lys	Arg	Pro	Met	Ala	Gly	Asp	Thr	Gly	Asn	
			355					360					365				
	Gly	Asn	Leu	Gln	Ala	Arg	Gly	Ala	Gly	Gly	Ser	Ser	Leu	Gly	Ile	Asp	
		370					375					380					
	Ala	Met	Met	Ala	Gly	Asp	Ala	Ile	Asn	Asn	Met	Ala	Leu	Gly	Lys	Leu	
	385					390					395					400	
25	Gly	Ala	Ala														

This hypersensitive response elicitor polypeptide or protein has a molecular weight of about 39 kDa, has a pI of approximately 4.3, and is heat stable at 100°C for at least 10 minutes. This hypersensitive response elicitor polypeptide or protein has substantially no cysteine. The hypersensitive response elicitor polypeptide or protein derived from *Erwinia amylovora* is more fully described in Wei, Z.-M., R. J. Laby, C. H. Zumoff, D. W. Bauer, S.-Y. He, A. Collmer, and S. V. Beer, "Harpin, Elicitor of the Hypersensitive Response Produced by the Plant Pathogen *Erwinia amylovora*," Science 257:85-88 (1992), which is hereby incorporated by reference. The DNA

- 12 -

molecule encoding this polypeptide or protein has a nucleotide sequence corresponding to SEQ. ID. No. 24 as follows:

	AAGCTTCGGC ATGGCACGTT TGACCGTTGG GTCGGCAGGG TACGTTTGAA TTATTCATAA	60
5	GAGGAATACG TTATGAGTCT GAATACAAGT GGGCTGGGAG CGTCAACGAT GCAAATTTCT	120
	ATCGGCGGTG CGGGCGGAAA TAACGGGTTG CTGGGTACCA GTCGCCAGAA TGCTGGGTTG	180
	GGTGGCAATT CTGCACTGGG GCTGGGCGGC GGTAAATCAA ATGATAACCGT CAATCAGCTG	240
	GCTGGCTTAC TCACCGGCAT GATGATGATG ATGAGCATCA TGGGCGGTGG TGGGCTGATG	300
	GGCGGTGGCT TAGGCGGTGG CTTAGGTAAT GGCTTGGGTG GCTCAGGTGG CCTGGGCGAA	360
10	GGACTGTCGA ACGCGCTGAA CGATATGTTA GCGGTTTCGC TGAACACGCT GGGCTCGAAA	420
	GGCGGCAACA ATACCACTTC AACAACAAAT TCCCCGCTGG ACCAGGCGCT GGGTATTAAAC	480
	TCAACGTCCC AAAACGACGA TTCCACCTCC GGCACAGATT CCACCTCAGA CTCCAGCGAC	540
	CCGATGCAGC AGCTGCTGAA GATGTTTCAGC GAGATAATGC AAAGCCTGTT TGGTGATGGG	600
	CAAGATGGCA CCCAGGGCAG TTCCTCTGGG GGCAAGCAGC CGACCGAAGG CGAGCAGAAC	660
15	GCCTATAAAA AAGGAGTCAC TGATGCGCTG TCGGGCCTGA TGGGTAATGG TCTGAGCCAG	720
	CTCCTTGGCA ACGGGGGACT GGGAGGTGGT CAGGGCGGTA ATGCTGGCAC GGGTCTTGAC	780
	GGTTCGTCGC TGGGCGGCAA AGGGCTGCAA AACCTGAGCG GGCCGGTGGA CTACCAGCAG	840
	TTAGGTAACG CCGTGGGTAC CGGTATCGGT ATGAAAGCGG GCATTTCAGGC GCTGAATGAT	900
	ATCGGTACGC ACAGGCACAG TTCAACCCGT TCTTTCGTCA ATAAAGGCGA TCGGGCGATG	960
20	GCGAAGGAAA TCGGTCAGTT CATGGACCAG TATCCTGAGG TGTTTGGCAA GCCGCAGTAC	1020
	CAGAAAGGCC CGGGTCAGGA GGTGAAAACC GATGACAAAT CATGGGCAAA AGCACTGAGC	1080
	AAGCCAGATG ACGACGGAAT GACACCAGCC AGTATGGAGC AGTTCAACAA AGCCAAGGGC	1140
	ATGATCAAAA GGCCCATGGC GGGTGATACC GGCAACGGCA ACCTGCAGGC ACGCGGTGCC	1200
	GGTGGTTCTT CGCTGGGTAT TGATGCCATG ATGGCCGGTG ATGCCATTAA CAATATGGCA	1260
25	CTTGGCAAGC TGGGCGCGGC TTAAGCTT	1288

The hypersensitive response elicitor polypeptide or protein derived from *Pseudomonas syringae* has an amino acid sequence corresponding to SEQ. ID. No. 25

as follows:

- 13 -

	Met	Gln	Ser	Leu	Ser	Leu	Asn	Ser	Ser	Ser	Leu	Gln	Thr	Pro	Ala	Met	
	1				5					10					15		
	Ala	Leu	Val	Leu	Val	Arg	Pro	Glu	Ala	Glu	Thr	Thr	Gly	Ser	Thr	Ser	
				20					25					30			
5	Ser	Lys	Ala	Leu	Gln	Glu	Val	Val	Val	Lys	Leu	Ala	Glu	Glu	Leu	Met	
			35					40					45				
	Arg	Asn	Gly	Gln	Leu	Asp	Asp	Ser	Ser	Pro	Leu	Gly	Lys	Leu	Leu	Ala	
		50					55					60					
10	Lys	Ser	Met	Ala	Ala	Asp	Gly	Lys	Ala	Gly	Gly	Gly	Ile	Glu	Asp	Val	
	65					70					75				80		
	Ile	Ala	Ala	Leu	Asp	Lys	Leu	Ile	His	Glu	Lys	Leu	Gly	Asp	Asn	Phe	
				85						90					95		
	Gly	Ala	Ser	Ala	Asp	Ser	Ala	Ser	Gly	Thr	Gly	Gln	Gln	Asp	Leu	Met	
				100					105					110			
15	Thr	Gln	Val	Leu	Asn	Gly	Leu	Ala	Lys	Ser	Met	Leu	Asp	Asp	Leu	Leu	
			115				120						125				
	Thr	Lys	Gln	Asp	Gly	Gly	Thr	Ser	Phe	Ser	Glu	Asp	Asp	Met	Pro	Met	
		130					135					140					
20	Leu	Asn	Lys	Ile	Ala	Gln	Phe	Met	Asp	Asp	Asn	Pro	Ala	Gln	Phe	Pro	
	145					150					155					160	
	Lys	Pro	Asp	Ser	Gly	Ser	Trp	Val	Asn	Glu	Leu	Lys	Glu	Asp	Asn	Phe	
				165						170					175		
	Leu	Asp	Gly	Asp	Glu	Thr	Ala	Ala	Phe	Arg	Ser	Ala	Leu	Asp	Ile	Ile	
			180						185					190			
25	Gly	Gln	Gln	Leu	Gly	Asn	Gln	Gln	Ser	Asp	Ala	Gly	Ser	Leu	Ala	Gly	
			195				200						205				
	Thr	Gly	Gly	Gly	Leu	Gly	Thr	Pro	Ser	Ser	Phe	Ser	Asn	Asn	Ser	Ser	
		210					215					220					
30	Val	Met	Gly	Asp	Pro	Leu	Ile	Asp	Ala	Asn	Thr	Gly	Pro	Gly	Asp	Ser	
	225					230					235					240	
	Gly	Asn	Thr	Arg	Gly	Glu	Ala	Gly	Gln	Leu	Ile	Gly	Glu	Leu	Ile	Asp	
				245					250						255		
	Arg	Gly	Leu	Gln	Ser	Val	Leu	Ala	Gly	Gly	Gly	Leu	Gly	Thr	Pro	Val	
			260						265					270			
35	Asn	Thr	Pro	Gln	Thr	Gly	Thr	Ser	Ala	Asn	Gly	Gly	Gln	Ser	Ala	Gln	
			275					280					285				

- 14 -

Asp Leu Asp Gln Leu Leu Gly Gly Leu Leu Leu Lys Gly Leu Glu Ala
 290 295 300
 Thr Leu Lys Asp Ala Gly Gln Thr Gly Thr Asp Val Gln Ser Ser Ala
 305 310 315 320
 5 Ala Gln Ile Ala Thr Leu Leu Val Ser Thr Leu Leu Gln Gly Thr Arg
 325 330 335
 Asn Gln Ala Ala Ala
 340

10

This hypersensitive response elicitor polypeptide or protein has a molecular weight of 34-35 kDa. It is rich in glycine (about 13.5%) and lacks cysteine and tyrosine.

Further information about the hypersensitive response elicitor derived from *Pseudomonas syringae* is found in He, S. Y., H. C. Huang, and A. Collmer,

15 "*Pseudomonas syringae* pv. *syringae* Harpin_{PS}: a Protein that is Secreted via the Hrp Pathway and Elicits the Hypersensitive Response in Plants," Cell 73:1255-1266 (1993), which is hereby incorporated by reference. The DNA molecule encoding the hypersensitive response elicitor from *Pseudomonas syringae* has a nucleotide sequence corresponding to SEQ. ID. No. 26 as follows:

20

ATGCAGAGTC TCAGTCTTAA CAGCAGCTCG CTGCAAACCC CGGCAATGGC CCTTGTCTCTG 60
 GTACGTCTCTG AAGCCGAGAC GACTGGCAGT ACGTCGAGCA AGGCGCTTCA GGAAGTTGTC 120
 GTGAAGCTGG CCGAGGAACT GATGCGCAAT GGTCAACTCG ACGACAGCTC GCCATTGGGA 180
 AAACGTGTTGG CCAAGTCGAT GGCCGCAGAT GGCAAGGCGG GCGGCGGTAT TGAGGATGTC 240
 25 ATCGCTGCGC TGGACAAGCT GATCCATGAA AAGCTCGGTG ACAACTTCGG CGCGTCTGCG 300
 GACAGCGCCT CGGGTACCGG ACAGCAGGAC CTGATGACTC AGGTGCTCAA TGGCCTGGCC 360
 AAGTCGATGC TCGATGATCT TCTGACCAAG CAGGATGGCG GGACAAGCTT CTCCGAAGAC 420
 GATATGCCGA TGCTGAACAA GATCGCGCAG TTCATGGATG ACAATCCCGC ACAGTTTCCC 480
 AAGCCGGACT CGGGCTCCTG GGTGAACGAA CTCAAGGAAG ACAACTTCCT TGATGGCGAC 540
 30 GAAACGGCTG CGTTCCGTTT GGCACCTCGAC ATCATTTGGCC AGCAACTGGG TAATCAGCAG 600
 AGTGACGCTG GCAGTCTGGC AGGGACGGGT GGAGGTCTGG GCACTCCGAG CAGTTTTTCC 660
 AACAACTCGT CCGTGATGGG TGATCCGCTG ATCGACGCCA ATACCGGTCC CGGTGACAGC 720
 GGCAATACCC GTGGTGAAGC GGGGCAACTG ATCGGCGAGC TTATCGACCG TGGCCTGCAA 780
 TCGGTATTGG CCGGTGGTGG ACTGGGCACA CCCGTAAACA CCCCAGAC CCGTACGTGC 840

- 15 -

GCGAATGGCG GACAGTCCGC TCAGGATCTT GATCAGTTGC TGGGCGGCTT GCTGCTCAAG 900
 GGCCTGGAGG CAACGCTCAA GGATGCCGGG CAAACAGGCA CCGACGTGCA GTCGAGCGCT 960
 GCGCAAATCG CCACCTTGCT GGTCACTACG CTGCTGCAAG GCACCCGCAA TCAGGCTGCA 1020
 GCCTGA 1026

5

The hypersensitive response elicitor polypeptide or protein derived from
Pseudomonas solanacearum has an amino acid sequence corresponding to SEQ. ID.
 No. 27 as follows:

10 Met Ser Val Gly Asn Ile Gln Ser Pro Ser Asn Leu Pro Gly Leu Gln
 1 5 10 15
 Asn Leu Asn Leu Asn Thr Asn Thr Asn Ser Gln Gln Ser Gly Gln Ser
 20 25 30
 15 Val Gln Asp Leu Ile Lys Gln Val Glu Lys Asp Ile Leu Asn Ile Ile
 35 40 45
 Ala Ala Leu Val Gln Lys Ala Ala Gln Ser Ala Gly Gly Asn Thr Gly
 50 55 60
 Asn Thr Gly Asn Ala Pro Ala Lys Asp Gly Asn Ala Asn Ala Gly Ala
 65 70 75 80
 20 Asn Asp Pro Ser Lys Asn Asp Pro Ser Lys Ser Gln Ala Pro Gln Ser
 85 90 95
 Ala Asn Lys Thr Gly Asn Val Asp Asp Ala Asn Asn Gln Asp Pro Met
 100 105 110
 25 Gln Ala Leu Met Gln Leu Leu Glu Asp Leu Val Lys Leu Leu Lys Ala
 115 120 125
 Ala Leu His Met Gln Gln Pro Gly Gly Asn Asp Lys Gly Asn Gly Val
 130 135 140
 Gly Gly Ala Asn Gly Ala Lys Gly Ala Gly Gly Gln Gly Gly Leu Ala
 145 150 155 160
 30 Glu Ala Leu Gln Glu Ile Glu Gln Ile Leu Ala Gln Leu Gly Gly Gly
 165 170 175
 Gly Ala Gly Ala Gly Gly Ala Gly Gly Gly Val Gly Gly Ala Gly Gly
 180 185 190
 35 Ala Asp Gly Gly Ser Gly Ala Gly Gly Ala Gly Gly Ala Asn Gly Ala
 195 200 205

- 16 -

	Asp Gly Gly Asn Gly Val Asn Gly Asn Gln Ala Asn Gly Pro Gln Asn	
	210 215 220	
	Ala Gly Asp Val Asn Gly Ala Asn Gly Ala Asp Asp Gly Ser Glu Asp	
	225 230 235 240	
5	Gln Gly Gly Leu Thr Gly Val Leu Gln Lys Leu Met Lys Ile Leu Asn	
	245 250 255	
	Ala Leu Val Gln Met Met Gln Gln Gly Gly Leu Gly Gly Gly Asn Gln	
	260 265 270	
10	Ala Gln Gly Gly Ser Lys Gly Ala Gly Asn Ala Ser Pro Ala Ser Gly	
	275 280 285	
	Ala Asn Pro Gly Ala Asn Gln Pro Gly Ser Ala Asp Asp Gln Ser Ser	
	290 295 300	
	Gly Gln Asn Asn Leu Gln Ser Gln Ile Met Asp Val Val Lys Glu Val	
	305 310 315 320	
15	Val Gln Ile Leu Gln Gln Met Leu Ala Ala Gln Asn Gly Gly Ser Gln	
	325 330 335	
	Gln Ser Thr Ser Thr Gln Pro Met	
	340	

It is encoded by a DNA molecule having a nucleotide sequence corresponding SEQ.

20 ID. No. 28 as follows:

	ATGTCAGTCG GAAACATCCA GAGCCCGTCG AACCTCCCGG GTCTGCAGAA CCTGAACCTC	60
	AACACCAACA CCAACAGCCA GCAATCGGGC CAGTCCGTGC AAGACCTGAT CAAGCAGGTC	120
	GAGAAGGACA TCCTCAACAT CATCGCAGCC CTCGTGCAGA AGGCCGCACA GTCGGCGGGC	180
	GGCAACACCG GTAACACCGG CAACGCGCCG GCGAAGGACG GCAATGCCAA CGCGGGCGCC	240
25	AACGACCCGA GCAAGAACGA CCCGAGCAAG AGCCAGGCTC CGCAGTCGGC CAACAAGACC	300
	GGCAACGTCG ACGACGCCAA CAACCAGGAT CCGATGCAAG CGCTGATGCA GCTGCTGGAA	360
	GACCTGGTGA AGCTGCTGAA GCGGGCCCTG CACATGCAGC AGCCCGGCGG CAATGACAAG	420
	GGCAACGGCG TGGGCGGTGC CAACGGCGCC AAGGGTGCCG GCGGCCAGGG CGGCCTGGCC	480
	GAAGCGCTGC AGGAGATCGA GCAGATCCTC GCCCAGCTCG GCGGCGGCGG TGCTGGCGCC	540
30	GGCGGCGCGG GTGGCGGTGT CGGCGGTGCT GGTGGCGCGG ATGGCGGCTC CGGTGCGGGT	600
	GGCGCAGGCG GTGCGAACGG CGCCGACGGC GGCAATGGCG TGAACGGCAA CCAGGCGAAC	660
	GGCCCGCAGA ACGCAGGCGA TGTCAACGGT GCCAACGGCG CGGATGACGG CAGCGAAGAC	720
	CAGGCGGGCC TCACCGGCGT GCTGCAAAAG CTGATGAAGA TCCTGAACGC GCTGGTGACG	780

- 17 -

```

ATGATGCAGC AAGGCGGCCT CGGCGGCGGC AACCAGGCGC AGGGCGGCTC GAAGGGTGCC      840
GGCAACGCCT CGCCGGCTTC CGGCGCGAAC CCGGGCGCGA ACCAGCCCGG TTCGGCGGAT      900
GATCAATCGT CCGGCCAGAA CAATCTGCAA TCCCAGATCA TGGATGTGGT GAAGGAGGTC      960
GTCCAGATCC TGCAGCAGAT GCTGGCGGCG CAGAACGGCG GCAGCCAGCA GTCCACCTCG     1020
5  ACGCAGCCGA TGTAAT                                         1035

```

Further information regarding the hypersensitive response elicitor polypeptide or protein derived from *Pseudomonas solanacearum* is set forth in Arlat, M., F. Van Gijsegem, J. C. Huet, J. C. Pemollet, and C. A. Boucher, "PopA1, a Protein which Induces a Hypersensitive-like Response in Specific Petunia Genotypes, is Secreted via the Hrp Pathway of *Pseudomonas solanacearum*," EMBO J. 13:543-533 (1994), which is hereby incorporated by reference.

The hypersensitive response elicitor polypeptide or protein from *Xanthomonas campestris* pv. *glycines* has an amino acid sequence corresponding to SEQ. ID. No. 29 as follows:

```

      Thr Leu Ile Glu Leu Met Ile Val Val Ala Ile Ile Ala Ile Leu Ala
      1           5           10           15
20  Ala Ile Ala Leu Pro Ala Tyr Gln Asp Tyr
      20           25

```

This sequence is an amino terminal sequence having only 26 residues from the hypersensitive response elicitor polypeptide or protein of *Xanthomonas campestris* pv. *glycines*. It matches with fimbrial subunit proteins determined in other *Xanthomonas campestris* pathovars.

The hypersensitive response elicitor polypeptide or protein from *Xanthomonas campestris* pv. *pelargonii* is heat stable, protease sensitive, and has a molecular weight of 20 kDa. It includes an amino acid sequence corresponding to SEQ. ID. No. 30 as follows:

```

      Ser Ser Gln Gln Ser Pro Ser Ala Gly Ser Glu Gln Gln Leu Asp Gln
      1           5           10           15
35  Leu Leu Ala Met
      20

```

- 18 -

Isolation of *Erwinia carotovora* hypersensitive response elicitor protein or polypeptide is described in Cui et al., "The RsmA Mutants of *Erwinia carotovora* subsp. *carotovora* Strain Ecc71 Overexpress *hrp* N_{Ecc} and Elicit a Hypersensitive Reaction-like Response in Tobacco Leaves," MPMI, 9(7):565-73 (1996), which is
5 hereby incorporated by reference. The hypersensitive response elicitor protein or polypeptide of *Erwinia stewartii* is set forth in Ahmad et al., "Harpin is Not Necessary for the Pathogenicity of *Erwinia stewartii* on Maize," 8th Int'l. Cong. Molec. Plant-Microbe Interact., July 14-19, 1996 and Ahmad, et al., "Harpin is Not Necessary for the Pathogenicity of *Erwinia stewartii* on Maize," Ann. Mtg. Am.
10 Phytopath. Soc., July 27-31, 1996, which are hereby incorporated by reference.

Hypersensitive response elicitor proteins or polypeptides from *Phytophthora parasitica*, *Phytophthora cryptogea*, *Phytophthora cinnamoni*, *Phytophthora capsici*, *Phytophthora megasperma*, and *Phytophthora citrophthora* are described in Kaman, et al., "Extracellular Protein Elicitors from Phytophthora: Most Specificity and
15 Induction of Resistance to Bacterial and Fungal Phytopathogens," Molec. Plant-Microbe Interact., 6(1):15-25 (1993), Ricci et al., "Structure and Activity of Proteins from Pathogenic Fungi Phytophthora Eliciting Necrosis and Acquired Resistance in Tobacco," Eur. J. Biochem., 183:555-63 (1989), Ricci et al., "Differential Production of Parasiticein, and Elicitor of Necrosis and Resistance in Tobacco, by Isolates of
20 *Phytophthora parasitica*," Plant Path., 41:298-307 (1992), Baillreul et al., "A New Elicitor of the Hypersensitive Response in Tobacco: A Fungal Glycoprotein Elicits Cell Death, Expression of Defence Genes, Production of Salicylic Acid, and Induction of Systemic Acquired Resistance," Plant J., 8(4):551-60 (1995), and Bonnet et al., "Acquired Resistance Triggered by Elicitors in Tobacco and Other Plants," Eur. J.
25 Plant Path., 102:181-92 (1996), which are hereby incorporated by reference.

The above elicitors are exemplary. Other elicitors can be identified by growing fungi or bacteria that elicit a hypersensitive response under which genes encoding an elicitor are expressed. Cell-free preparations from culture supernatants can be tested for elicitor activity (i.e. local necrosis) by using them to infiltrate
30 appropriate plant tissues.

- 19 -

Fragments of the above hypersensitive response elicitor polypeptides or proteins as well as fragments of full length elicitors from other pathogens are encompassed by the method of the present invention.

Suitable fragments can be produced by several means. In the first, subclones
5 of the gene encoding a known elicitor protein are produced by conventional molecular genetic manipulation by subcloning gene fragments. The subclones then are expressed *in vitro* or *in vivo* in bacterial cells to yield a smaller protein or peptide that can be tested for elicitor activity according to the procedure described below.

As an alternative, fragments of an elicitor protein can be produced by
10 digestion of a full-length elicitor protein with proteolytic enzymes like chymotrypsin or *Staphylococcus* proteinase A, or trypsin. Different proteolytic enzymes are likely to cleave elicitor proteins at different sites based on the amino acid sequence of the elicitor protein. Some of the fragments that result from proteolysis may be active elicitors of resistance.

15 In another approach, based on knowledge of the primary structure of the protein, fragments of the elicitor protein gene may be synthesized by using the PCR technique together with specific sets of primers chosen to represent particular portions of the protein. These then would be cloned into an appropriate vector for expression of a truncated peptide or protein.

20 Chemical synthesis can also be used to make suitable fragments. Such a synthesis is carried out using known amino acid sequences for the elicitor being produced. Alternatively, subjecting a full length elicitor to high temperatures and pressures will produce fragments. These fragments can then be separated by conventional procedures (e.g., chromatography, SDS-PAGE).

25 An example of suitable fragments of an *Erwinia* hypersensitive response elicitor which fragments elicit a hypersensitive response are fragments of the *Erwinia amylovora* hypersensitive response elicitor. Suitable fragments include a C-terminal fragment of the amino acid sequence of SEQ. ID. No. 23, an N-terminal fragment of the amino acid sequence of SEQ. ID. No. 23, or an internal fragment of the amino
30 acid sequence of SEQ. ID. No. 23. The C-terminal fragment of the amino acid sequence of SEQ. ID. No. 23 can span amino acids 105 and 403 of SEQ. ID. No. 23. The N-terminal fragment of the amino acid sequence of SEQ. ID. No. 23 can span the

- 20 -

following amino acids of SEQ. ID. No. 23: 1 and 98, 1 and 104, 1 and 122, 1 and 168, 1 and 218, 1 and 266, 1 and 342, 1 and 321, and 1 and 372. The internal fragment of the amino acid sequence of SEQ. ID. No. 23 can span the following amino acids of SEQ. ID. No. 23: 76 and 209, 105 and 209, 99 and 209, 137 and 204, 137 and 200, 109 and 204, 109 and 200, 137 and 180, and 105 and 180. Other suitable fragments can be identified in accordance with the present invention.

Variants may be made by, for example, the deletion or addition of amino acids that have minimal influence on the properties, secondary structure and hydropathic nature of the polypeptide. For example, a polypeptide may be conjugated to a signal (or leader) sequence at the N-terminal end of the protein which co-translationally or post-translationally directs transfer of the protein. The polypeptide may also be conjugated to a linker or other sequence for ease of synthesis, purification, or identification of the polypeptide.

The fragment of the present invention is preferably produced in purified form (preferably at least about 60%, more preferably 80%, pure) by conventional techniques. Typically, the fragment of the present invention is produced but not secreted into the growth medium of recombinant host cells. Alternatively, the protein or polypeptide of the present invention is secreted into growth medium. In the case of unsecreted protein, to isolate the protein fragment, the host cell (e.g., *E. coli*) carrying a recombinant plasmid is propagated, lysed by sonication, heat, or chemical treatment, and the homogenate is centrifuged to remove bacterial debris. The supernatant is then subjected to heat treatment and the fragment is separated by centrifugation. The supernatant fraction containing the fragment is subjected to gel filtration in an appropriately sized dextran or polyacrylamide column to separate the fragment. If necessary, the protein fraction may be further purified by ion exchange or HPLC.

The DNA molecule encoding the fragment of the hypersensitive response elicitor polypeptide or protein can be incorporated in cells using conventional recombinant DNA technology. Generally, this involves inserting the DNA molecule into an expression system to which the DNA molecule is heterologous (i.e. not normally present). The heterologous DNA molecule is inserted into the expression system or vector in proper sense orientation and correct reading frame. The vector

- 21 -

contains the necessary elements for the transcription and translation of the inserted protein-coding sequences.

U.S. Patent No. 4,237,224 to Cohen and Boyer, which is hereby incorporated by reference, describes the production of expression systems in the form of recombinant plasmids using restriction enzyme cleavage and ligation with DNA ligase. These recombinant plasmids are then introduced by means of transformation and replicated in unicellular cultures including procaryotic organisms and eucaryotic cells grown in tissue culture.

Recombinant genes may also be introduced into viruses, such as vaccina virus. Recombinant viruses can be generated by transection of plasmids into cells infected with virus.

Suitable vectors include, but are not limited to, the following viral vectors such as lambda vector system gt11, gt WES.tB, Charon 4, and plasmid vectors such as pBR322, pBR325, pACYC177, pACYC1084, pUC8, pUC9, pUC18, pUC19, pLG339, pR290, pKC37, pKC101, SV 40, pBluescript II SK +/- or KS +/- (see "Stratagene Cloning Systems" Catalog (1993) from Stratagene, La Jolla, Calif, which is hereby incorporated by reference), pQE, pIH821, pGEX, pET series (see F.W. Studier et. al., "Use of T7 RNA Polymerase to Direct Expression of Cloned Genes," Gene Expression Technology vol. 185 (1990), which is hereby incorporated by reference), and any derivatives thereof. Recombinant molecules can be introduced into cells via transformation, particularly transduction, conjugation, mobilization, or electroporation. The DNA sequences are cloned into the vector using standard cloning procedures in the art, as described by Sambrook et al., Molecular Cloning: A Laboratory Manual, Cold Springs Laboratory, Cold Springs Harbor, New York (1989), which is hereby incorporated by reference.

A variety of host-vector systems may be utilized to express the protein-encoding sequence(s). Primarily, the vector system must be compatible with the host cell used. Host-vector systems include but are not limited to the following: bacteria transformed with bacteriophage DNA, plasmid DNA, or cosmid DNA; microorganisms such as yeast containing yeast vectors; mammalian cell systems infected with virus (e.g., vaccinia virus, adenovirus, etc.); insect cell systems infected with virus (e.g., baculovirus); and plant cells infected by bacteria. The expression

- 22 -

elements of these vectors vary in their strength and specificities. Depending upon the host-vector system utilized, any one of a number of suitable transcription and translation elements can be used.

5 Different genetic signals and processing events control many levels of gene expression (e.g., DNA transcription and messenger RNA (mRNA) translation).

Transcription of DNA is dependent upon the presence of a promotor which is a DNA sequence that directs the binding of RNA polymerase and thereby promotes mRNA synthesis. The DNA sequences of eucaryotic promotors differ from those of procaryotic promotors. Furthermore, eucaryotic promotors and accompanying genetic
10 signals may not be recognized in or may not function in a procaryotic system, and, further, procaryotic promotors are not recognized and do not function in eucaryotic cells.

Similarly, translation of mRNA in procaryotes depends upon the presence of the proper procaryotic signals which differ from those of eucaryotes. Efficient
15 translation of mRNA in procaryotes requires a ribosome binding site called the Shine-Dalgarno ("SD") sequence on the mRNA. This sequence is a short nucleotide sequence of mRNA that is located before the start codon, usually AUG, which encodes the amino-terminal methionine of the protein. The SD sequences are complementary to the 3'-end of the 16S rRNA (ribosomal RNA) and probably
20 promote binding of mRNA to ribosomes by duplexing with the rRNA to allow correct positioning of the ribosome. For a review on maximizing gene expression, see Roberts and Lauer, Methods in Enzymology, 68:473 (1979), which is hereby incorporated by reference.

Promotors vary in their "strength" (i.e. their ability to promote transcription).
25 For the purposes of expressing a cloned gene, it is desirable to use strong promotors in order to obtain a high level of transcription and, hence, expression of the gene. Depending upon the host cell system utilized, any one of a number of suitable promotors may be used. For instance, when cloning in *E. coli*, its bacteriophages, or plasmids, promotors such as the T7 phage promotor, *lac* promotor, *trp* promotor, *recA*
30 promotor, ribosomal RNA promotor, the P_R and P_L promotors of coliphage lambda and others, including but not limited, to *lacUV5*, *ompF*, *bla*, *lpp*, and the like, may be used to direct high levels of transcription of adjacent DNA segments. Additionally, a

- 23 -

hybrid *trp-lacUV5 (tac)* promotor or other *E. coli* promotors produced by recombinant DNA or other synthetic DNA techniques may be used to provide for transcription of the inserted gene.

5 Bacterial host cell strains and expression vectors may be chosen which inhibit the action of the promotor unless specifically induced. In certain operations, the addition of specific inducers is necessary for efficient transcription of the inserted DNA. For example, the *lac* operon is induced by the addition of lactose or IPTG (isopropylthio-beta-D-galactoside). A variety of other operons, such as *trp*, *pro*, etc., are under different controls.

10 Specific initiation signals are also required for efficient gene transcription and translation in procaryotic cells. These transcription and translation initiation signals may vary in "strength" as measured by the quantity of gene specific messenger RNA and protein synthesized, respectively. The DNA expression vector, which contains a promotor, may also contain any combination of various "strong" transcription and/or
15 translation initiation signals. For instance, efficient translation in *E. coli* requires an SD sequence about 7-9 bases 5' to the initiation codon ("ATG") to provide a ribosome binding site. Thus, any SD-ATG combination that can be utilized by host cell ribosomes may be employed. Such combinations include but are not limited to the SD-ATG combination from the *cro* gene or the *N* gene of coliphage lambda, or
20 from the *E. coli* tryptophan E, D, C, B or A genes. Additionally, any SD-ATG combination produced by recombinant DNA or other techniques involving incorporation of synthetic nucleotides may be used.

Once the isolated DNA molecule encoding the fragment of a hypersensitive response elicitor polypeptide or protein has been cloned into an expression system, it
25 is ready to be incorporated into a host cell. Such incorporation can be carried out by the various forms of transformation noted above, depending upon the vector/host cell system. Suitable host cells include, but are not limited to, bacteria, virus, yeast, mammalian cells, insect, plant, and the like.

The present invention further relates to methods of imparting disease
30 resistance to plants, enhancing plant growth, and/or effecting insect control for plants. These methods involve applying the fragment of a hypersensitive response elicitor polypeptide or protein, which fragment itself elicits a hypersensitive response, in a

- 24 -

non-infectious form to all or part of a plant or a plant seed under conditions effective for the fragment to impart disease resistance, enhance growth, and/or control insects. Alternatively, these fragments of a hypersensitive response elicitor protein or polypeptide can be applied to plants such that seeds recovered from such plants
5 themselves are able to impart disease resistance in plants, to enhance plant growth, and/or to effect insect control.

As an alternative to applying a fragment of a hypersensitive response elicitor polypeptide or protein to plants or plant seeds in order to impart disease resistance in plants, to effect plant growth, and/or to control insects on the plants or plants grown
10 from the seeds, transgenic plants or plant seeds can be utilized. When utilizing transgenic plants, this involves providing a transgenic plant transformed with a DNA molecule encoding a fragment of a hypersensitive response elicitor polypeptide or protein, which fragment elicits a hypersensitive response, and growing the plant under conditions effective to permit that DNA molecule to impart disease resistance to
15 plants, to enhance plant growth, and/or to control insects. Alternatively, a transgenic plant seed transformed with a DNA molecule encoding a fragment of a hypersensitive response elicitor polypeptide or protein which fragment elicits a hypersensitive response can be provided and planted in soil. A plant is then propagated from the planted seed under conditions effective to permit that DNA molecule to impart
20 disease resistance to plants, to enhance plant growth, and/or to control insects.

The embodiment of the present invention where the hypersensitive response elicitor polypeptide or protein is applied to the plant or plant seed can be carried out in a number of ways, including: 1) application of an isolated fragment or 2) application of bacteria which do not cause disease and are transformed with a genes encoding the
25 fragment. In the latter embodiment, the fragment can be applied to plants or plant seeds by applying bacteria containing the DNA molecule encoding the fragment of the hypersensitive response elicitor polypeptide or protein which fragment elicits a hypersensitive response. Such bacteria must be capable of secreting or exporting the fragment so that the fragment can contact plant or plant seeds cells. In these
30 embodiments, the fragment is produced by the bacteria *in planta* or on seeds or just prior to introduction of the bacteria to the plants or plant seeds.

- 25 -

The methods of the present invention can be utilized to treat a wide variety of plants or their seeds to impart disease resistance, enhance growth, and/or control insects. Suitable plants include dicots and monocots. More particularly, useful crop plants can include: alfalfa, rice, wheat, barley, rye, cotton, sunflower, peanut, corn, 5 potato, sweet potato, bean, pea, chicory, lettuce, endive, cabbage, brussel sprout, beet, parsnip, turnip, cauliflower, broccoli, turnip, radish, spinach, onion, garlic, eggplant, pepper, celery, carrot, squash, pumpkin, zucchini, cucumber, apple, pear, melon, citrus, strawberry, grape, raspberry, pineapple, soybean, tobacco, tomato, sorghum, and sugarcane. Examples of suitable ornamental plants are: *Arabidopsis thaliana*, 10 *Saintpaulia*, petunia, pelargonium, poinsettia, chrysanthemum, carnation, and zinnia.

With regard to the use of the fragments of the hypersensitive response elicitor protein or polypeptide of the present invention in imparting disease resistance, absolute immunity against infection may not be conferred, but the severity of the disease is reduced and symptom development is delayed. Lesion number, lesion size, 15 and extent of sporulation of fungal pathogens are all decreased. This method of imparting disease resistance has the potential for treating previously untreatable diseases, treating diseases systemically which might not be treated separately due to cost, and avoiding the use of infectious agents or environmentally harmful materials.

The method of imparting pathogen resistance to plants in accordance with the 20 present invention is useful in imparting resistance to a wide variety of pathogens including viruses, bacteria, and fungi. Resistance, *inter alia*, to the following viruses can be achieved by the method of the present invention: *Tobacco mosaic virus* and *Tomato mosaic virus*. Resistance, *inter alia*, to the following bacteria can also be imparted to plants in accordance with present invention: *Pseudomonas solanacearum*, 25 *Pseudomonas syringae* pv. *tabaci*, and *Xanthomonas campestris* pv. *pelargonii*. Plants can be made resistant, *inter alia*, to the following fungi by use of the method of the present invention: *Fusarium oxysporum* and *Phytophthora infestans*.

With regard to the use of the fragments of the hypersensitive response elicitor protein or polypeptide of the present invention to enhance plant growth, various forms 30 of plant growth enhancement or promotion can be achieved. This can occur as early as when plant growth begins from seeds or later in the life of a plant. For example, plant growth according to the present invention encompasses greater yield, increased

- 26 -

quantity of seeds produced, increased percentage of seeds germinated, increased plant size, greater biomass, more and bigger fruit, earlier fruit coloration, and earlier fruit and plant maturation. As a result, the present invention provides significant economic benefit to growers. For example, early germination and early maturation permit crops
5 to be grown in areas where short growing seasons would otherwise preclude their growth in that locale. Increased percentage of seed germination results in improved crop stands and more efficient seed use. Greater yield, increased size, and enhanced biomass production allow greater revenue generation from a given plot of land.

Another aspect of the present invention is directed to effecting any form of
10 insect control for plants. For example, insect control according to the present invention encompasses preventing insects from contacting plants to which the hypersensitive response elicitor has been applied, preventing direct insect damage to plants by feeding injury, causing insects to depart from such plants, killing insects proximate to such plants, interfering with insect larval feeding on such plants,
15 preventing insects from colonizing host plants, preventing colonizing insects from releasing phytotoxins, etc. The present invention also prevents subsequent disease damage to plants resulting from insect infection.

The present invention is effective against a wide variety of insects. European corn borer is a major pest of corn (dent and sweet corn) but also feeds on over 200
20 plant species including green, wax, and lima beans and edible soybeans, peppers, potato, and tomato plus many weed species. Additional insect larval feeding pests which damage a wide variety of vegetable crops include the following: beet armyworm, cabbage looper, corn ear worm, fall armyworm, diamondback moth, cabbage root maggot, onion maggot, seed corn maggot, pickleworm (melonworm),
25 pepper maggot, tomato pinworm, and maggots. Collectively, this group of insect pests represents the most economically important group of pests for vegetable production worldwide.

The method of the present invention involving application of the fragment of a hypersensitive response elicitor polypeptide or protein, which fragment elicits a
30 hypersensitive response, can be carried out through a variety of procedures when all or part of the plant is treated, including leaves, stems, roots, etc. This may (but need not) involve infiltration of the hypersensitive response elicitor polypeptide or protein

- 27 -

into the plant. Suitable application methods include high or low pressure spraying, injection, and leaf abrasion proximate to when elicitor application takes place. When treating plant seeds or propagules (e.g., cuttings), in accordance with the application embodiment of the present invention, the fragment of the hypersensitive response elicitor protein or polypeptide, in accordance with present invention, can be applied by low or high pressure spraying, coating, immersion, or injection. Other suitable application procedures can be envisioned by those skilled in the art provided they are able to effect contact of the fragment with cells of the plant or plant seed. Once treated with the fragment of the hypersensitive response elicitor of the present invention, the seeds can be planted in natural or artificial soil and cultivated using conventional procedures to produce plants. After plants have been propagated from seeds treated in accordance with the present invention, the plants may be treated with one or more applications of the fragment of the hypersensitive response elicitor protein or polypeptide or whole elicitors to impart disease resistance to plants, to enhance plant growth, and/or to control insects on the plants.

The fragment of the hypersensitive response elicitor polypeptide or protein, in accordance with the present invention, can be applied to plants or plant seeds alone or in a mixture with other materials. Alternatively, the fragment can be applied separately to plants with other materials being applied at different times.

A composition suitable for treating plants or plant seeds in accordance with the application embodiment of the present invention contains a fragment of a hypersensitive response elicitor polypeptide or protein which fragment elicits a hypersensitive response in a carrier. Suitable carriers include water, aqueous solutions, slurries, or dry powders. In this embodiment, the composition contains greater than 500 nM of the fragment.

Although not required, this composition may contain additional additives including fertilizer, insecticide, fungicide, nematocide, and mixtures thereof. Suitable fertilizers include $(\text{NH}_4)_2\text{NO}_3$. An example of a suitable insecticide is Malathion. Useful fungicides include Captan.

Other suitable additives include buffering agents, wetting agents, coating agents, and abrading agents. These materials can be used to facilitate the process of the present invention. In addition, the hypersensitive response eliciting fragment can

- 28 -

be applied to plant seeds with other conventional seed formulation and treatment materials, including clays and polysaccharides.

In the alternative embodiment of the present invention involving the use of transgenic plants and transgenic seeds, a hypersensitive response eliciting fragment
5 need not be applied topically to the plants or seeds. Instead, transgenic plants transformed with a DNA molecule encoding such a fragment are produced according to procedures well known in the art.

The vector described above can be microinjected directly into plant cells by use of micropipettes to transfer mechanically the recombinant DNA. Crossway, Mol. Gen. Genetics, 202:179-85 (1985), which is hereby incorporated by reference. The
10 genetic material may also be transferred into the plant cell using polyethylene glycol. Krens, et al., Nature, 296:72-74 (1982), which is hereby incorporated by reference.

Another approach to transforming plant cells with a gene which imparts resistance to pathogens is particle bombardment (also known as biolistic
15 transformation) of the host cell. This can be accomplished in one of several ways. The first involves propelling inert or biologically active particles at cells. This technique is disclosed in U.S. Patent Nos. 4,945,050, 5,036,006, and 5,100,792, all to Sanford et al., which are hereby incorporated by reference. Generally, this procedure involves propelling inert or biologically active particles at the cells under conditions
20 effective to penetrate the outer surface of the cell and to be incorporated within the interior thereof. When inert particles are utilized, the vector can be introduced into the cell by coating the particles with the vector containing the heterologous DNA. Alternatively, the target cell can be surrounded by the vector so that the vector is carried into the cell by the wake of the particle. Biologically active particles (e.g.,
25 dried bacterial cells containing the vector and heterologous DNA) can also be propelled into plant cells.

Yet another method of introduction is fusion of protoplasts with other entities, either minicells, cells, lysosomes, or other fusible lipid-surfaced bodies. Fraley, et al., Proc. Natl. Acad. Sci. USA, 79:1859-63 (1982), which is hereby incorporated by
30 reference.

The DNA molecule may also be introduced into the plant cells by electroporation. Fromm et al., Proc. Natl. Acad. Sci. USA, 82:5824 (1985), which is

- 29 -

hereby incorporated by reference. In this technique, plant protoplasts are electroporated in the presence of plasmids containing the expression cassette. Electrical impulses of high field strength reversibly permeabilize biomembranes allowing the introduction of the plasmids. Electroporated plant protoplasts reform the cell wall, divide, and regenerate.

Another method of introducing the DNA molecule into plant cells is to infect a plant cell with *Agrobacterium tumefaciens* or *A. rhizogenes* previously transformed with the gene. Under appropriate conditions known in the art, the transformed plant cells are grown to form shoots or roots, and develop further into plants. Generally, this procedure involves inoculating the plant tissue with a suspension of bacteria and incubating the tissue for 48 to 72 hours on regeneration medium without antibiotics at 25-28°C.

Agrobacterium is a representative genus of the gram-negative family Rhizobiaceae. Its species are responsible for crown gall (*A. tumefaciens*) and hairy root disease (*A. rhizogenes*). The plant cells in crown gall tumors and hairy roots are induced to produce amino acid derivatives known as opines, which are catabolized only by the bacteria. The bacterial genes responsible for expression of opines are a convenient source of control elements for chimeric expression cassettes. In addition, assaying for the presence of opines can be used to identify transformed tissue.

Heterologous genetic sequences can be introduced into appropriate plant cells, by means of the Ti plasmid of *A. tumefaciens* or the Ri plasmid of *A. rhizogenes*. The Ti or Ri plasmid is transmitted to plant cells on infection by *Agrobacterium* and is stably integrated into the plant genome. J. Schell, Science, 237:1176-83 (1987), which is hereby incorporated by reference.

After transformation, the transformed plant cells must be regenerated.

Plant regeneration from cultured protoplasts is described in Evans et al., Handbook of Plant Cell Cultures, Vol. 1: (MacMillan Publishing Co., New York, 1983); and Vasil I.R. (ed.), Cell Culture and Somatic Cell Genetics of Plants, Acad. Press, Orlando, Vol. I, 1984, and Vol. III (1986), which are hereby incorporated by reference.

- 30 -

It is known that practically all plants can be regenerated from cultured cells or tissues, including but not limited to, all major species of sugarcane, sugar beets, cotton, fruit trees, and legumes.

Means for regeneration vary from species to species of plants, but generally a suspension of transformed protoplasts or a petri plate containing transformed explants is first provided. Callus tissue is formed and shoots may be induced from callus and subsequently rooted. Alternatively, embryo formation can be induced in the callus tissue. These embryos germinate as natural embryos to form plants. The culture media will generally contain various amino acids and hormones, such as auxin and cytokinins. It is also advantageous to add glutamic acid and proline to the medium, especially for such species as corn and alfalfa. Efficient regeneration will depend on the medium, on the genotype, and on the history of the culture. If these three variables are controlled, then regeneration is usually reproducible and repeatable.

After the expression cassette is stably incorporated in transgenic plants, it can be transferred to other plants by sexual crossing. Any of a number of standard breeding techniques can be used, depending upon the species to be crossed.

Once transgenic plants of this type are produced, the plants themselves can be cultivated in accordance with conventional procedure with the presence of the gene encoding the hypersensitive response eliciting fragment resulting in disease resistance, enhanced plant growth, and/or control of insects on the plant. Alternatively, transgenic seeds or propagules (e.g., cuttings) are recovered from the transgenic plants. The seeds can then be planted in the soil and cultivated using conventional procedures to produce transgenic plants. The transgenic plants are propagated from the planted transgenic seeds under conditions effective to impart disease resistance to plants, to enhance plant growth, and/or to control insects. While not wishing to be bound by theory, such disease resistance, growth enhancement, and/or insect control may be RNA mediated or may result from expression of the polypeptide or protein fragment.

When transgenic plants and plant seeds are used in accordance with the present invention, they additionally can be treated with the same materials as are used to treat the plants and seeds to which a hypersensitive response eliciting fragment is applied. These other materials, including hypersensitive response eliciting fragments,

- 31 -

- can be applied to the transgenic plants and plant seeds by the above-noted procedures, including high or low pressure spraying, injection, coating, and immersion. Similarly, after plants have been propagated from the transgenic plant seeds, the plants may be treated with one or more applications of the hypersensitive response eliciting
- 5 fragment to impart disease resistance, enhance growth, and/or control insects. Such plants may also be treated with conventional plant treatment agents (e.g., insecticides, fertilizers, etc.).

EXAMPLES

10

Example 1 - Strains and plasmids used

The strains and plasmids used are set forth in Table 1 below

15

Table 1

Plasmid name	<i>E. amylovora</i> source strain	Brief Description, Relevant Phenotype, Reference	Harpin, fragment (or NA)
pBCKS	-	Cm ^r derivative of pBluescript KS. Stratagene, La Jolla, CA	
pBCSK	-	Cm ^r derivative of pBluescript SK. Stratagene, La Jolla, CA	
pBSKS	-	pBluescript KS; Ap ^r . Stratagene, La Jolla, CA	
pBSSK II	-	pBluescript II SK; Ap ^r . Stratagene, La Jolla, CA	
pBW7	-	Mob ⁺ Tc ^r helper plasmid. (Rella, et al., "Transposon Insertion Mutagenesis of <i>Pseudomonas aeruginosa</i> With a Tn5 Derivative: Application to Physical Mapping of the <i>arc</i> Gene Cluster," <i>Gene</i> 33:293-303 (1985), which is hereby incorporated by reference)	NA
pCPP51	-	Broad host range derivative of pBSSK II containing <i>ori</i> from pRO1614.	NA
PCPP430	321	Functional <i>hrp</i> gene cluster of <i>E. amylovora</i> Ea321 cloned into pCPP9. Beer, S.V., et al., "The <i>hrp</i> Gene Cluster of <i>Erwinia Amylovora</i> ," In Hennecke, H., and D.P.S. Verma, (eds.), <i>Advances in Molecular Genetics of Plant-Microbe Interactions</i> , Kluwer Academic Publishers, Dordrecht, Netherlands, 53-60 (1991), which is hereby incorporated by reference.	
pCPP460	246	Functional <i>hrp</i> gene cluster of <i>E. amylovora</i> Ea246 cloned into pCPP9.	
pCPP1104	321	1.2 kb <i>Pst</i> I fragment of pCPP1084 in pBCKS	Ea C31
pCPP1105	321	1.1 kb <i>Sma</i> I fragment of pCPP1084 in pBCSK	Ea C82
pCPp1107	321	1.3 kb <i>Hind</i> III fragment of pCPP1084 in pBCSK	Ea wt
pCPp1108	321	1.2 kb <i>Hinc</i> II- <i>Hind</i> III of pCPP1084 in pBCSK	Ea N11
pCPP1109	321	pCPP1107 with internal <i>Ava</i> II fragment deleted	Ea 1175
pCPP1110	321	As pCPP1108, but cloned into pBCKS	Ea N9
pCPP1111	321	367 bp <i>Taq</i> I fragment of pCPP1107 in pBCSK	Ea C305

pCPP1113	246	As pCPP1109, but 425bp <i>Avall</i> fragment of pCPP1098 deleted	Ear 1175
pCPP1119	246	Site specific mutation in pCPP1098; stop codon inserted at L36	Ear C368
pCPP1120	246	Site specific mutation in pCPP1098; stop codon inserted at T123	Ear C281
pCPP1121	321	702bp <i>KpnI</i> fragment internal to <i>hrpN</i> deleted	Ea C375
pCPP1127	246	3.1 kb <i>BamHI</i> fragment of pCPP1098 in pSU21	Ear wt
pCPP1128	246	Tn10 minikan in pCPP1127	Ear undef
pCPP1136	246	4.4 kb <i>EcoRI</i> fragment of pCPP1120, religated	EAR N122
pCPP1146	246	4.2 kb <i>EcoRI</i> fragment of pCPP1119, religated	Ear N35
pCPP1147	321	1.2 kb <i>BamHI</i> fragment of pCPP1084, PCR amplified, cloned into pSU23	Ea wt
pCPP1148	246	As pCpP1147, but from pCPP1098	Ear wt
pCPP1150	246	As pCpP1148, but in pCPP51 vector	Ear wt
pCPP1163	246	3.1 kb <i>BamHI</i> fragment of pCPP1098 in pCPP51	Ear wt
pCPP1164	321	1.3 kb <i>HindIII</i> of pCpP1084 in pCPP51	Ea wt
pCPP1165		Derivative of pCPP51 w/ <i>KpnI</i> , <i>SacII</i> sites removed	NA
pCPP1167	321	1.3 kb <i>HindII</i> fragment of pCpP1107 in pCPP1165	Ea wt
pCPP1169	246	As pCPP1167, but 3.1 kb <i>BamHI</i> insert from pCpP1098	Ear wt
pCPP1170	246	PCPP1098; Σ -Sp ligated into <i>EcoRV</i> site	Ear C128 Σ
pCPP1171	246	<i>KpnI</i> fragment internal to <i>hrpN</i> deleted; shifted reading frame	Ea C375
pCPP1172	321	Derivative of pCPP1167 with in-frame deletion of <i>KpnI</i> fragment internal to <i>hrpN</i>	Ea 1235
pCPP1173	246	As pCPP1172, but from pCPP1169	Ear 1235
pCPP217	321	PCPP1084 with 2 <i>StyI</i> fragments deleted, blunted, and religated	Ea C185
pCPP1252	321	PCPP1105 with Σ -Sp ligated at <i>SmaI</i> site	Ea C82 Σ
pGPI-2		P15a ori.T7 RNA polymerase-encoding plasmid; for protein expression (Tabor, S., et al., "A Bacteriophage T7 DNA Polymerase/Promoter System for Controlled Exclusive Expression of Specific Genes," <u>Proc. Natl. Acad. Sci. USA</u> , 82:1074-1078 (1985), which is hereby incorporated by reference).	N/A
PHP45 Σ Ω		Ap ^r ; Sp ^r ; source of Ω -Sp fragment; (Fellay, R., et al., "Interposon Mutagenesis of Soil and Water Bacteria a Family of DNA Fragments Designed for <i>in vitro</i> Insertional Mutagenesis of Gram-Negative Bacteria," <u>Gene</u> , 52:147-154 (1987), which is hereby incorporated by reference).	N/A
pSU21		P15a ori Km ^r (Bartolomé, B.Y., et al., "Construction and properties of a Family of pACYC184-Derived Cloning Vectors Compatible With pBR322 and its Derivatives," <u>Gene</u> , 102:75-78 (1991), which is hereby incorporated by reference).	N/A
PSU23		P15a ori Km ^r (Bartolomé, B.Y., et al., "Construction and properties of a Family of pACYC184-Derived Cloning Vectors Compatible With pBR322 and its Derivatives," <u>Gene</u> , 102:75-78 (1991), which is hereby incorporated by reference).	N/A
Strains used			
<i>E. amylovora</i>		Ea273Nx; Nalidixic acid resistant (Nx ^r) derivative of Ea273. CUCPB 2348	
<i>E. amylovora</i>		Rifampin resistant derivative of Ea32. CUCPB 2545	

- 33 -

<i>E. coli</i>	GM272; <i>dam</i> ⁻ , <i>dcm</i> ⁻ . CUCPB 3047; (Blumenthal, R.M., et al., " <i>E. coli</i> Can Restrict Methylated DNA and May Skew Genomic Libraries," <u>Trends in Biotech.</u> , 4:302-305 (1986), which is hereby incorporated by reference)
<i>E. coli</i>	BL21(DE3); CUCPB 4277; (Studier, F.W., and B.A. Moffatt, "Use of Bacteriophage T7 RNA Polymerase to Direct Selective High-level Expression of Cloned Genes," <u>J. Mol. Biol.</u> , 189:113-130 (1986), which is hereby incorporated by reference)
<i>E. coli</i>	DH5 α ; (N \times ⁺). CUCPB 2475; Stratagene, La Jolla, CA.

Example 2 - Molecular biology techniques.

- 5 Several approaches were employed to obtain truncated or otherwise altered versions of both *E. amylovora* harpins. These techniques included: (i) subcloning of restriction fragments containing portions of the gene encoding the hypersensitive response elicitor protein or polypeptide from *Erwinia amylovora* (i.e. *hrpN*) into expression vectors, by standard techniques (Sambrook, et al., Molecular Cloning: a
- 10 Laboratory Manual, 2nd ed. ed. Cold Spring Harbor, Laboratory," Cold Spring Harbor, NY (1989), which is hereby incorporated by reference); (ii) insertion of an Ω -fragment (Fellay, et al., "Interposon Mutagenesis of Soil and Water Bacteria a Family of DNA Fragments Designed for *in vitro* Insertional Mutagenesis of Gram-Negative Bacteria," Gene 52:147-154 (1987), which is hereby incorporated by reference) into
- 15 *hrpN*; (iii) site-specific mutagenesis approaches (Innis, et al., PCR Protocols. A Guide to Methods and Applications, Academic Press San Diego, CA (1990); Kunkel, et al., "Rapid and Efficient Site-Specific Mutagenesis Without Phenotypic Selection," Proc. Nat. Acad. Sci. USA 82:488-492 (1985), which are hereby incorporated by reference); and (iv) creation of nested deletions (Erase-a-BaseTM kit; Promega,
- 20 Madison, WI). C-terminal deletion analysis of the hypersensitive response elicitor protein or polypeptide from *Erwinia amylovora* (i.e. harpin_{Ea}) in pCPP1084 could not be performed because of the location of restriction enzyme cleavage sites in pCPP1084. For N-terminal deletions, pCPP1084 DNA was prepared using a Qiagen midiprep column (Qiagen, Chatsworth, CA) and digested with *sst I* followed by
- 25 *EcoRI*. Subsequently, the digested DNA was subjected to exonuclease III digestion, ligation, and transformation into *E. coli* BL21(DE3). Deletion sizes were estimated by agarose gel electrophoresis. Harpin fragments were named with respect to the portion of harpin deleted (e.g., harpin_{Ea} C82 lacks the C-terminal 82 amino acid residues of full-length harpin_{Ea}).

- 34 -

Example 3 - Protein expression.

For expression from T7 promoters, T7 RNA polymerase-dependent systems were used. These systems utilized either strain *E. coli* BL21(DE3) (Studier, et al.,
5 "Use of Bacteriophage T7 RNA Polymerase to Direct Selective High-Level Expression of Cloned Gene," J. Mol. Biol. 189:113-130 (1986), which is hereby incorporated by reference), or plasmid pGP1-2 (Tabor, et al., "A Bacteriophage T7 DNA Polymerase/Promoter System for Controlled Exclusive Expression of Specific Genes," Proc. Natl. Acad. Sci., USA 82:1074-1078 (1985), which is hereby
10 incorporated by reference) in *E. coli* DH5 α . Expression of *hrpN* from the T7 promoter was induced by addition of IPTG to a final concentration of 0.4 mM. For expression in *E. amylovora* Ea321 (i.e. harpin_{Ea}) or Ea273, pGP1-2 was introduced by transformation with a 42°C heat shock for 10 minutes, or by electroporation (Biorad Gene PulserTM). Hypersensitive response (i.e. HR)-eliciting activity was screened in
15 tobacco cv. Xanthi leaves by *in planta* lysis (He, et al., "*Pseudomonas syringae* pv. *syringae* harpin_{Pss}: a Protein That is Secreted via the Hrp Pathway and Elicits the Hypersensitive Response in Plants," Cell 73:1255-1266 (1993), which is hereby incorporated by reference) or by preparation of boiled and unboiled "CFEPs" (Wei, et al., "Harpin, Elicitor of the Hypersensitive Response Produced by the Plant Pathogen
20 *Erwinia amylovora*," Science 257:85-88 (1992), which is hereby incorporated by reference).

Example 4 - In vitro proteolysis of harpin.

25 *In vitro* proteolysis of harpin_{Ea} with *Staphylococcus* V8 proteinase (also termed endoproteinase Glu-C), trypsin, pepsin, and papain was performed as recommended (Scopes, et al., Protein Purification: Principles and Practice, 2nd ed. Springer-Verlag. New York (1987), which is hereby incorporated by reference), for 2-
16 hrs. at 20-37°. Endoproteinase Glu-C digestion was performed either in 50 mM
30 ammonium bicarbonate, pH 7.8 (in which cleavage occurs only after glutamic acid), or in 50 mM potassium phosphate, pH 7.8 (in which cleavage after both glutamic acid and aspartic acid occurs).

- 35 -

Example 5 - Plant-derived proteinases.

Intercellular fluids (IF) were obtained from tobacco, tomato, apple, raspberry, and cotoneaster, as described (Hammond-Kosack, et al., "Preparation and Analysis of
5 Intercellular Fluid," p. 15-21. *In* S.J. Gurr, M.J. McPherson, and D.J. Bowles (ed.), Molecular Plant Pathology A Practical Approach, 2nd ed., The Practical Approach Series, IRL Publishers, Oxford (1992), which is hereby incorporated by reference), by vacuum infiltration of intercellular spaces with high-purity water. Proteolytic digestion of PAGE-purified harpin_{EA} was performed for 2-16 hrs. at 20-37°C, pH, by
10 mixing equal volumes of IF with harpin_{EA}. A total leaf extract was obtained by grinding tobacco leaf panels with mortar and pestle in 5 mM potassium phosphate. The extract was centrifuged and filtered, and the clarified ground leaf extract used identically as was the IF. Proteinase inhibitors were employed as follows: Pepstatin A (final concentration 1µM), E-64 (1µM), Aprotinin (2µg/ml), o-phenanthroline
15 (1mM), and p-mercuribenzoate (PCMB) (Sigma, St. Louis, MO).

Example 6 - Peptide purification.

Peptide fragments of harpin obtained following digestion with tobacco IF were
20 fractionated by reverse-phase HPLC on a Vydac C18 column using a 2-60% acetonitrile gradient in 0.1% trifluoroacetic acid. Fractions were lyophilized, resuspended in 5 mM potassium phosphate and infiltrated into tobacco leaf panels. The fraction with greatest HR-eliciting activity was refractionated as above with a 35-70% acetonitrile gradient, and the purity of each fraction was assayed via gas
25 chromatography-mass spectroscopy (GC-MS) and by N-terminal protein sequencing at the Cornell Biotechnology Program Core Facility.

Example 7 - Proteinase activity-stained gels.

30 Proteinase activity of IF was assayed in activity-stained polyacrylamide gels (Laemmli, "Cleavage of Structural Proteins During the Assembly of the Head of Bacteriophage T4," Nature 227:680-685 (1970), which is hereby incorporated by reference) copolymerized with 0.1% gelatin (Heussen, et al., "Electrophoretic Analysis of Plasminogen Activators in Polyacrylamide Gels Containing Sodium

- 36 -

Dodecyl Sulfate and Copolymerized Substrates," Anal. Biochem. 102:196-202 (1980), which is hereby incorporated by reference). After electrophoresis, each gel was rinsed extensively to remove SDS and allow refolding of proteinases in the gel. Following additional incubation to allow proteolysis to occur, the gels were stained
5 with 0.1% Amido Black in 30% methanol/10% acetic acid. Each gel stained darkly (due to the presence of copolymerized gelatin) except where proteinases had digested the gelatin, resulting in colorless bands representing the sites of proteinase activity.

Example 8 - Truncated harpins retain HR-eliciting activity.

10

The stability and the HR-eliciting activity of proteins encoded by various DNA constructs is shown in Figure 1. Many DNA constructs encoding portions of harpin_{Ea} or harpin_{Ea}r did not yield detectable protein products following induction of expression in the T7 promoter-polymerase system (Tabor, et al., "A Bacteriophage T7
15 DNA Polymerase/Promoter System for Controlled Exclusive Expression of Specific Genes," Proc. Natl. Acad. Sci. USA 82:1074-1078 (1985), which is hereby incorporated by reference) and analysis of cell extracts by PAGE, possibly due to instability of the encoded proteins. No DNA constructs (e.g., those obtained via Erase-a-BaseTM protocol) yielded detectable protein products displaying N-terminal
20 deletions relative to the full-length protein. No stable but inactive proteins were identified. Several constructs encoding proteins truncated at their C-terminus and often including additional vector-encoded amino acids yielded detectable products (e.g. harpin_{Ea} C82). In contrast, a construct encoding the same 321 N-terminal amino acid residues of harpin_{Ea}, but yielding a protein truncated by the presence of an Ω-
25 fragment (harpin_{Ea} C82 Ω) was unstable (i.e. no product was detected). A construct encoding a harpin_{Ea} fragment with a large internal deletion (harpin_{Ea} I175) was also successfully used to express protein. These various truncated proteins were tested for HR-eliciting activity. A 98 residue N-terminal harpin_{Ea} fragment (harpin_{Ea} C305) was the smallest bacterially-produced peptide that displayed HR-eliciting activity.

30

- 37 -

Example 9 - Secretion of harpin_{Ea} with an altered C-terminus.

The effect of alteration at the harpin C-terminus on its secretion was examined. Harpin C31 contains the N-terminal 372 amino acids of harpin, but lacks
5 the C-terminal 31 residues, which are replaced by 47 residues encoded by the vector, resulting in a protein slightly larger than the wild type harpin_{Ea}. The C31 protein retains HR-eliciting activity and is stable and easily expressed and detected by western analysis or PAGE but it is no longer secreted into the culture supernatant as is the wild type protein (Figure 2). The presence of harpin_{Ea} C31 does not interfere with
10 secretion of the wild type harpin, which is found in both the CFEP and the culture supernatant. However, harpin_{Ea} C31 is found only in the CFEP.

Example 10 - Effect of proteolysis on harpin_{Ea}'s HR eliciting activity

15 In order to generate additional harpin_{Ea} fragments, purified full length protein was proteolyzed *in vitro* by several proteinases, including endoproteinase Glu-C, trypsin, pepsin, and papain (e.g., Figures 3 and 4). Harpin solutions digested with trypsin or with papain lost all activity. In contrast, following digestion with endoproteinase Glu-C, HR-eliciting activity was retained. No peptides larger than 6
20 kD were evident by PAGE following trypsin digestion. Endoproteinase Glu-C digestion yielded an approximately 20 kD fragment, larger than expected if all cleavage sites were cut, indicating that digestion was not complete (Figure 4).

Example 11 - Apoplastic fluids (IF) contain harpin-degrading proteolytic activity

25 Apoplastic fluids (intercellular fluids; IF) from tobacco and other plants were also employed to proteolyze harpin. Each IF tested possessed proteinase activity(s), as indicated by the presence of multiple activity-stained bands in polyacrylamide gels containing co-polymerized gelatin (Figures 5A to 5C), as well as by the disappearance
30 of detectable harpin_{Ea} (Schägger, et al., "Tricine-Sodium Dodecyl Sulfate Gel Electrophoresis for the Separation of Proteins in the Range From 1 to 100 kDa," Anal. Biochem. 166:368-379 (1987), which is hereby incorporated by reference) following overnight digestion of purified harpin_{Ea} with IF. Proteinase activity was substantially greater at 37°C than at 20°C, and activity was higher at pH 8.5 than at pH 7. Several

- 38 -

inhibitors were employed in order to define the proteolytic activity(s) of the IF. No single proteinase inhibitor which was employed prevented degradation of harpin_{Ea}. However, a mixture of the inhibitors Pepstatin A (1 μ M), E-64 (1 μ M), Aprotinin (2 μ g/ml), and o-phenanthroline (1 mM), targeted at acid proteinases, cysteine
5 proteinases, serine proteinases, and metalloproteinases, respectively, partially inhibited proteolysis.

Harpin_{Ea} degraded by proteolytic activities present in the plant apoplast retained HR-eliciting activity (Figure 3). In contrast, harpin_{Ea} proteolyzed by a clarified extract produced by grinding tobacco leaf tissue with mortar and pestle lost
10 HR-eliciting activity. In order to study whether apoplastic degradation of harpin was a prerequisite to its HR-eliciting activity, the length of time required for leaf collapse when either intact harpin or harpin predigested with tobacco IF was infiltrated into tobacco leaf panels was compared. Both preparations elicited the HR in a similar time frame (12-18 hours, depending on the experiment).

15

Example 12 - Characterization of HR-eliciting peptide fragments

Peptides resulting from digestion by apoplastic plant proteinase(s) were fractionated by reverse phase HPLC (Vydac C18 column), and tested for activity.
20 Following treatment of intact harpin_{Ea} with tobacco IF, three fractions contained some HR-eliciting activity on tobacco. Two of the three demonstrated weak activity, and little protein was present. They were not further characterized. Fraction 19, which contained the strongest activity as well as the most protein, was refractionated using a more shallow elution gradient (Figure 6). Refractionation, N-terminal protein
25 sequencing, and molecular weight analysis by mass spectroscopy indicated that four largely overlapping peptides were present. Peak 19-1 contained peptides P91 and P95, corresponding to harpin_{Ea} residues 110-200 and 110-204; peak 19-2 contained peptides P64 and P68, corresponding to harpin_{Ea} residues 137-200 and 137-204. 19-1 and 19-2 each possessed HR-eliciting activity. The smallest peptide thus confirmed
30 to retain activity consisted of residues 137-204. The two peptides in each peak were not separable under the conditions used. These active fragments are distinct from the smallest active N-terminal fragment (harpin_{Ea}C305), and indicate that more than one portion of harpin_{Ea} displays activity *in planta*. Further digestion with trypsin

- 39 -

abolished the HR-eliciting activity of 19-2. This proteinase cleaves P64 and P68 as shown in Figure 7. Further digestion with endoproteinase Glu-C in ammonium bicarbonate buffer abolished the HR-eliciting activity of 19-1. Endoproteinase Glu-C is predicted to cleave P91 and P95 as shown in Figure 7. Loss of elicitor-activity followed further digestion of these peptides with endoproteinase Glu-C or trypsin.

Example 13 - *E. amylovora* harpin's similarity with other proteins.

The predicted protein sequences of proteinaceous HR elicitors from several other bacterial plant pathogens, and of other proteins known to be, or thought to be, secreted by a type III secretion pathway were also compared with that of harpin_{Ea}. When harpin_{Ea} was compared with elicitors from *E. amylovora* Ea246 (i.e. harpin_{Ea}), *Erwinia chrysanthemi* EC16 (harpin_{Ech}) (Bauer, et al., "Erwinia chrysanthemi harpin_{Ech}: An Elicitor of the Hypersensitive Response That Contributes to Soft-Rot Pathogenesis," Mol. Plant-Microbe Interact 8:484-491 (1995), which is hereby incorporated by reference), *Erwinia carotovora* subsp. *carotovora* (harpin_{Ecc}) (Mukherjee, et al., Presented at the 8th International Congress Molecular Plant-Microbe Interactions, Knoxville, TN (1996), which is hereby incorporated by reference), *Erwinia stewartii* (Harpin_{Es}) (Frederick, et al., "The wts Water-Soaking Genes of *Erwinia stewartii* are Related to *hrp* genes," Presented at the Seventh International Symposium on Molecular Plant-Microbe Interactions, Edinburgh, Scotland (1994), which is hereby incorporated by reference), *Ralstonia* (*Pseudomonas*) *solanacearum* (PopA) (Arlat, et al., "PopA1, a Protein Which Induces a Hypersensitivity-Like Response on Specific *Petunia* Genotypes, is Secreted via the Hrp Pathway of *Pseudomonas solanacearum*," EMBO J. 13:543-553 (1994), which is hereby incorporated by reference), *Pseudomonas syringae* 61 (harpin_{Pss}) (He, et al., "Pseudomonas syringae pv. syringae harpin_{Pss}: a Protein That is Secreted via the Hrp Pathway and Elicits the Hypersensitive Response in Plants," Cell 73:1255-1266 (1993), which is hereby incorporated by reference), *Pseudomonas syringae* pv. *tomato* (harpin_{Pst}) (Preston, et al., "The HrpZ Proteins of *Pseudomonas syringae* pvs. *syringae*, *glycinea*, and *tomato* Are Encoded By An Operon Containing *Yersinia ysc* Homologs and Elicit the Hypersensitive Response in Tomato But Not Soybean," Mol. Plant-Microbe Interact 8:717-732 (1995), which is hereby incorporated by reference),

- 40 -

the *Erwinia*-derived harpins contained significant regions of similarity at the C-terminus. In addition, all the elicitors are glycine-rich, secreted, and heat-stable. Limited similarity between harpin_{PSS} and harpin_{Ea} had been reported previously (He, et al., "*Pseudomonas syringae* pv. *syringae* harpin_{PSS}: a Protein That is Secreted via the Hrp Pathway and Elicits the Hypersensitive Response in Plants," Cell 73:1255-1266 (1993), which is hereby incorporated by reference), (Laby, et al., Presented at the Seventh International Symposium on Molecular Plant-Microbe Interactions, Edinburgh, Scotland (1994), which is hereby incorporated by reference). A limited region of similarity between harpin_{Ea} and other harpins from *Erwinia* spp. was also evident at the extreme N-terminus of each protein, where 9 out of the first 26 residues are conserved (Figure 8). Kyte-Doolittle hydrophobicity plots for each of the harpins produced by different *Erwinia* spp. are illustrated in Figure 9. Each of the *Erwinia* harpins examined displays a generally similar hydrophobicity profile along the full length of the protein. This profile is distinct from the profile demonstrated by PopA1 and by harpin_{PSS}, and does not possess the symmetry evident in the profile of those two proteins. The hydrophobicity profile of each *Erwinia* harpin is generally similar to that of the others, yet distinct from that reported for harpin_{PSS} (Alfano, et al., "Analysis of the Role of the *Pseudomonas Syringae* HrpZ harpin in Elicitation of the Hypersensitive Response to Tobacco Using Functionally Nonpolar *hrpZ* Deletion Mutations, Truncated HrpZ Fragments, and *hrmA* Mutations," Mol. Microbiol. 19:715-728 (1996), which is hereby incorporated by reference). Harpin_{Ecc} possesses a strikingly hydrophobic region around residues 54-143 (Mukherjee, et al., Presented at the 8th International Congress Molecular Plant-Microbe Interactions, Knoxville, TN (1996), which is hereby incorporated by reference). This portion of the protein is also the most hydrophobic region of harpin_{Ea} and harpin_{Es}. The rest of each protein is predominantly hydrophilic.

Truncated proteins and fragments of harpin obtained following proteolytic digestion of the full length protein indicate several surprising aspects of harpin_{Ea} HR-eliciting activity. These harpin fragments demonstrate that HR-eliciting activity resides in distinct regions of the protein, and that relatively small fragments of the protein, as little as 68 residues and possibly less, are sufficient for this activity. Fragments of other plant pathogen-derived elicitor proteins also retain biological

- 41 -

activity, including Avr9 from *Cladosporium fulvum* (Van den Ackerveken, et al., "The AVR9 Race-Specific Elicitor of *Cladosporium fulvum* is Processed by Endogenous and Plant Proteases," Pl. Physiol. 103:91-96 (1993), which is hereby incorporated by reference), Pep-13 of *Phytophthora megasperma* (Nürnberg, et al.,
5 "High Affinity Binding of a Fungal Oligopeptide Elicitor to the Parsley Plasma Membranes Triggers Multiple Defence Responses," Cell, 78:449-460 (1994), which is hereby incorporated by reference), and harpin_{Pss} of *P. syringae* pv. *syringae* (Alfano, et al., "Analysis of the Role of the *Pseudomonas syringae* HrpZ harpin in Elicitation of the Hypersensitive Response in Tobacco Using Functionally Nonpolar *hrpZ*
10 Deletion Mutations, Truncated HrpZ Fragments, and *hrmA* Mutations," Mol. Microbiol. 19:715-728 (1996), which is hereby incorporated by reference).

Expression of truncated harpin fragments and proteolysis of full-length harpins showed that two distinct fragments retain HR-eliciting activity. The primary sequence of each active fragment show no discernable similarity with each other, or
15 with other elicitor-active peptides. However, the sites of cleavage by trypsin and endoproteinase Glu-C suggest portions of each fragment required for activity. It would be of interest to alter specifically the amino acid residues at or near these cleavage sites to determine whether HR-eliciting activity is altered or lost. Additionally, harpin_{Ea} P64 and P68 demonstrate distinct hydrophobicity during
20 reverse-phase HPLC (Figure 6), and they correspond to a hydrophobic peak in a Kyte-Doolittle plot (Figure 9). The role of this putative hydrophobic domain could be tested by mutagenesis, or by synthesis of altered peptides. However, the fact that multiple fragments independently possess HR-eliciting activity complicates analysis of full-length proteins.

25 This finding, that fragments of the protein retain HR-eliciting activity, also allowed (at least) two apoplastic proteinase activities, which are distinct from intracellular plant proteinases, to be defined. Two apoplastic plant proteinases (from soybean) have been studied in some detail. SMEP, a metalloproteinase (Huangpu, et al., "Purification and Developmental Analysis of an Extracellular Proteinase From
30 Young Leaves of Soybean," Plant Physiol 108:969-974 (1995); McGeehan, et al., "Sequencing and Characterization of the Soybean Leaf Metalloproteinase," Plant Physiol. 99:1179-1183 (1992), which are hereby incorporated by reference) sensitive

- 42 -

to EDTA, is thought to cleave at G/L and G/I. Interestingly, although there are 19 potential SMEP cleavage sites in the intact harpin_{Ea}, only one of them is located within fragments P91 and P95, and none are within fragments P64 and P68 (Figure 7). P91 and P95 thus may represent partial digestion products of a SMEP-like proteinase in the tobacco apoplast. The other studies soybean apoplastic proteinase, SLAP, a

5 sulfhydryl proteinase (Huangpu, et al., "Purification and Developmental Analysis of an Extracellular Proteinase From Young Leaves of Soybean," Plant Physiol. 108:969-974 (1995), which is hereby incorporated by reference) sensitive to p-chloromercuribenzoic acid (pCMB). Several lines of evidence suggest that multiple

10 proteolytic activities in the IF are degrading harpin_{Ea}. PMSF, a serine protease inhibitor, decreases but does not entirely block harpin_{Ea} degradation (Figure 5C); no single proteinase inhibitor tested blocks harpin degradation, and the cleavage sites after residues 109, 136, 200, and 204 are dissimilar. Endoproteinase Glu-C does not abolish activity of full-length harpin, but does abolish activity of P91 and P95 (and

15 presumably P64 and P68); trypsin abolishes the activity of P64 and P68 (and presumably P91 and P95). These final digests suggest specific portions of each distinct HR-eliciting peptide which are presumably necessary for its activity, as mentioned previously.

The apoplastic activities degrade harpin without destroying its HR-eliciting

20 ability, in contrast to intracellular proteolytic activities present in a ground leaf-extract, which abolish activity. This raises a number of intriguing questions, e.g., does the plant use these harpin fragments as elicitor-signals? The timing of the HR was examined when full length harpin and harpin predigested by tobacco intercellular fluid were each infiltrated into tobacco leaves. The HR elicited by each preparation

25 occurred 12-18 hours after infiltration. Co-infiltration of proteinase inhibitors into tobacco leaf panels along with harpin also had no effect on harpin's HR eliciting activity, although limited proteolytic degradation cannot be ruled out in this case, particularly since it seems that at least two, and perhaps more, apoplastic proteinases are present in tobacco. Because predigested harpin elicited the HR no faster than

30 undigested protein, proteolytic digestion seems unlikely to be a rate-limiting step required for HR to occur. The role of these apoplastic proteinases which are able to hydrolyze harpin partially, yet unable to abolish harpin's HR-eliciting activity on

- 43 -

tobacco, remains unclear. Salzer et al., "Rapid Reactions of Spruce Cells to Elicitors Released From the Ectomycorrhizal Fungus *Hebeloma crustuliniforme* and Inactivation of These Elicitors by Extracellular Spruce Cell Enzymes," Planta 198:118-126 (1996), which is hereby incorporated by reference, have noted that

5 spruce (*Picea abies* (L.) Karst.) modulates the level of fungal cell wall elicitors released by the ectomycorrhizal fungus *Hebeloma crustuliniforme* by inactivating these molecules in the apoplast. They propose that *Picea* controls the elicitor level as part of its symbiotic interaction with the fungus. Similarly, PGIP of *Phaseolus vulgaris* has been suggested to modulate the level of elicitor-active

10 oligogalacturonides present during the plant-parasite interaction in bean (Desiderio, et al., "Polygalacturonase, PGIP, and Oligogalacturonides in Cell-Cell Communication," Biochem. Sci. Trans. 22:394-397 (1994), which is hereby incorporated by reference). Perhaps the retention of HR-eliciting activity by harpin fragments contributes to the ability of plants to recognize the presence of a pathogen. In this regard, it would be

15 interesting to explore whether transgenic host and non-host plants, engineered for apoplastic expression of a harpin activity-degrading proteinase, would exhibit reduced or increased sensitivity to *E. amylovora*, compared to non-engineered plants.

Despite numerous attempts, only a handful of truncated derivatives of harpin_{Ea} and harpin_{Ea}r were successfully expressed from portions of *hrpN*. Problems with

20 protein stability were evident in that several truncated harpins were unstable and difficult to purify. Additionally, expression of truncated harpins may be deleterious to bacteria. Truncated harpin_{Ea}C31 was, however, stable and easily purified, but not secreted, suggesting that C-terminal sequences are involved in harpin secretion. Unfortunately, the presence of vector-encoded amino acids in this protein complicates

25 this conclusion. All attempts to clone β -galactosidase-harpin fusion proteins have been unsuccessful, as were attempts to clone and express *hrpN* downstream of the *lacZ* promoter in several commonly used vectors such as pBluescript. Expression of such constructs is evidently deleterious to bacterial strains.

Wei, et al., "Harpin, Elicitor of the Hypersensitive Response Produced By the

30 Plant Pathogen *Erwinia amylovora*," Science, 257:85-88 (1992), which is hereby incorporated by reference, reported previously that BLAST searches indicated harpin_{Ea} possessed slight similarity to several other glycine-rich proteins, including

- 44 -

keratins and glycine-rich cell wall proteins. However, this is thought to be due to the high glycine content of harpin_{EA}, and as such does not suggest a role for harpin_{EA}. Examination of N-terminal sequences from several HR-eliciting proteins produced by phytopathogenic bacteria (Figure 8) yielded some potential similarities. However, the region in question is quite short. The region of putative primary sequence similarity is limited to the first 26 residues at the N-terminus, and its role remains unclear. Surprisingly, *E. carotovora* harpin_{Ecc} appears more similar to the harpins from *E. amylovora* and *E. stewartii* than to that from *E. chrysanthemi*, to which it is more closely related with respect to its taxonomic position as well as its mechanism of pathogenicity (i.e. soft-rots). In addition, although primary sequence similarity is strongest only in the C-terminal third of each protein, the *Erwinia* harpins possess broadly similar hydrophobicity profiles along their entire lengths (Figure 9). Based on its hydrophobicity profile, Alfano, et al., "Analysis of the Role of the *Pseudomonas syringae* HrpZ harpin in Elicitation of the Hypersensitive Response in Tobacco Using Functionally Nonpolar *hrpZ* Deletion Mutations, Truncated HrpZ Fragments, and *hrmA* Mutations," Mol. Microbiol. 19:715-728 (1996), which is hereby incorporated by reference, speculated that harpin_{Pss} may have an amphiphilic nature. However, the *Erwinia* harpins' profiles do not match that of harpin_{Pss}.

Recently, a number of other secreted glycine-rich pathogenicity associated proteins, elicitors of the HR or other plant-defense responses, have been described from other plant pathogenic bacteria and fungi (Boller, "Chemoperception of Microbial Signals in Plant Cells," Ann. Rev. Plant Physiol. Plant Molec. Biol. 46:189-214 (1996), which is hereby incorporated by reference), including *Phytophthora megasperma* (Ballieul, et al., "A New Elicitor of the Hypersensitive Response in Tobacco: a Fungal Glycoprotein Elicits Cell Death, Expression of Defence Genes, Production of Salicylic Acid, and Induction of Systemic Acquired Resistance," Plant Journal 8:551-560 (1995); Nürnberger, et al., "High Affinity Binding of a Fungal Oligopeptide Elicitor to the Parsley Plasma Membranes Triggers Multiple Defence Responses," Cell 78:449-460 (1994), which are hereby incorporated by reference), and *Magnaporthe grisea* (Sweigard, et al., "Identification, Cloning, and Characterization of PWL2, a Gene For Host Species Specificity in the Rice Blast Fungus," Plant Cell 7:1221-1233 (1995), which is hereby incorporated by reference).

- 45 -

Proteinaceous HR-elicitors have also now been described from *Phyechosporium secalis* (Rohe, et al., "The Race-Specific Elicitor, NIP1, From the Barley Pathogen, *Rhynchosporium secalis*, Determines Avirulence on Host Plants of the *Rrs1* Resistance Genotype," EMBO Journal 14:4168-4177 (1995) which is hereby
5 incorporated by reference, while *P. infestans* (Pieterse, et al., "Structure and Genomic Organization of the *ipiB* and *ipiO* Gene Clusters of *Phytophthora infestans*," Gene, 138:67-77 (1994), which is hereby incorporated by reference) produces a glycine-rich pathogenicity-associated family of proteins of unknown function. Because the
10 primary amino acid sequence of each elicitor protein or peptide fragment shows no obvious similarity to that of the others, it is unclear whether they interact with the same target on or in the plant cell, plasma membrane, or cell wall. In that regard, it might be of interest to test whether any one of these molecules inhibits the action of other(s). The increasing availability of peptides such as Pep13, Avr9, P68, and harpin_{Ea} C305 with plant-defense response-eliciting activity (HR and otherwise)
15 should enable precise probing of their targets on or in plant cells, as well as determination of whether their mechanisms of activity are similar, distinct, or overlapping.

20 Example 14 - Bacterial strains and plasmids

Escherichia coli stains used in the following examples include DH5 α and BL21(DE3) purchased from Gibco BRL and Stratagene, respectively. The pET28(b) vector was purchased from Novagen. Eco DH5 α /2139 contained the complete hrpN gene. The 2139 construct was produced by D. Bauer at the Cornell University. The
25 hrpN gene was cleaved from the 2139 plasmid by restriction enzyme digestion with HindIII, then purified from an agarose gel to serve as the DNA template for PCR synthesis of truncated hrpN clones. These clones were subsequently inserted into the (His)₆ vector pET28(b) which contained a Kan^r gene for selection of transformants.

30 Example 15 - DNA Manipulation

Restriction enzymes were obtained from Boehringer Mannheim or Gibco BRL. T4 DNA ligase, Calf Intestinal Alkaline Phosphatase (CIAP), and PCR

- 46 -

Supermix™ were obtained from Gibco BRL. The QIAprep Spin Miniprep Kit, the Qiagen Plasmid Mini Kit, and the QIAquick PCR Purification Kit were purchased from Qiagen. The PCR primers were synthesized by Lofstrand Labs Limited (Gaithersburg, MD). The oligopeptides were synthesized by Bio-Synthesis, Inc. (Lewisville, TX). All DNA manipulations such as plasmid isolation, restriction enzyme digestion, DNA ligation, and PCR were performed according to standard techniques (molecular cloning) or protocols provided by the manufacturer.

Example 16 - Fragmentation of *hrpN* gene

10 A series of N-terminal and C-terminal truncated *hrpN* genes and internal fragments were generated via PCR (Fig. 10). The full length *hrpN* gene was used as the DNA template and 3' and 5' primers were designed for each truncated clone (Fig. 11). The 3' primers contained in the NdeI enzyme cutting site which contained
15 the start codon ATG (Methionine) and the 5' primers contained the stop codon TAA and a HindIII enzyme cutting site for ligation into the pET28(b) vector. PCR was carried out in 0.5 ml tubes in a GeneAmp™ 9600 or 9700. 45 µl of Supermix™ were mixed with 20 pmoles of each pair of DNA primers, 10 ng of full length harpin DNA, and diH₂O to a final volume of 50 µl. After heating the mixture at 95°C for 2 min, the
20 PCR was performed for 30 cycles at 94°C for 1 min, 58°C for 1 min and 72°C for 1.5 min. The PCR products were verified on a 6% TBE gel (Novex). Amplified DNA was purified with the QIAquick PCR purification kit, digested with Nde I and Hind III at 37°C for 5 hours, extracted once with phenol:chloroform:isoamylalcohol (25:25:1) and precipitated with ethanol. 5 µg of pET28(b) vector DNA were digested
25 with 15 units of Nde I and 20 units of Hind III at 37°C for 3 hours followed with CIAP treatment to reduce the background resulting from incomplete single enzyme digestion. Digested vector DNA was purified with the QIAquick PCR purification kit and directly used for ligation. Ligation was carried out at 14-16°C for 5-12 hours in a 15 µl mixture containing ca. 200 ng of digested pET28(b), 30 ng of targeted PCR
30 fragment, and 1 unit T4 DNA ligase. 5 - 7.5 µl of ligation solution were added to 100 µl of DH5α competent cells in a 15 ml falcon tube and incubated on ice for 30 min. After a heat shock at 42°C for 45 seconds, 0.9 ml SOC solution or 0.45 ml LB media

- 47 -

were added to each tube and incubated at 37°C for 1 hour. 20, 100, and 200 µl of transformed cells were placed onto LB agar with 30 µg/ml of kanamycin and incubated at 37°C overnight. Single colonies were transferred to 3 ml LB-media and incubated overnight at 37°C. Plasmid DNA was prepared from 2 ml of culture with the QIAprep Miniprep kit. The DNA from the transformed cells was analyzed by restriction enzyme digestion or partial sequencing to verify the success of the transformations. Plasmids with the desired DNA sequence were transferred into the BL21 strain using the standard chemical transformation method as indicated above. A clone containing the full length harpin protein in the pet28(b) vector was generated as a positive control, and a clone with only the pET28(b) vector was generated as a negative control.

Example 17 - Expression of harpin truncated proteins

Escherichia coli BL21(DE3) strains containing the hrpN clones were grown in Luria broth medium (g/L Difco Yeast extract, 10 g/L Difco Tryptone, 5g/L NaCl, and 1 mM NaOH) containing 30 µg/ml of kanamycin at 37°C overnight. The bacteria were then inoculated into 100 volumes of the same medium and grown at 37°C to an OD₆₂₀ of 0.6-0.8. The bacteria were then inoculated into 250 volumes of the same medium and grown at 37°C to an OD₆₂₀ of ca. 0.3 or 0.6-0.8. One milli molar IPTG was then added and the cultures grown at 19°C overnight (ca. 18 hours). Not all of the clones were successfully expressed using this strategy. Several of the clones had to be grown in Terrific broth (12 g/L Bacto Tryptone, 24 g/L Bacto yeast, 0.4% glycerol, 0.17 M KH₂PO₄, and 0.72 K₂HPO₄), and/or grown at 37°C after IPTG induction, and/or harvested earlier than overnight (Table 2).

- 48 -

Table 2: Expression of harpin truncated proteins

Fragment	amino acids (SEQ. ID. No. 23)	Growth medium	Induction O.D.	Expression temp.	Harvest time
1 (+ control)	1-403	LB	ca. 0.3 or 0.6-0.8	19°C or 25°C	16-18 hr
2 (+ control)	-	LB and TB	ca. 0.3 or 0.6-0.8	19 C and 37 C	16-18 hr
3	105-403	LB	0.6-0.8	19°C	16-18 hr
4	169-403	TB	ca. 0.3	19°C	16-18 hr
5	210-403	LB or M9ZB	0.6-0.8	19°C	16-18 hr
6	257-403	LB or M9ZB	0.6-0.8	19°C	16-18 hr
7	343-403	LB	ca. 0.3	19°C	5 hr
8	1-75	TB	ca. 0.3	37°C	16-18 hr
9	1-104	TB	ca. 0.3	37°C	16-18 hr
10	1-168	TB	ca. 0.3	37°C	16-18 hr
11	1-266	LB	ca. 0.3	37°C	4 hr
12	1-342	LB	0.6-0.8	19°C	16-18 hr
13	76-209	LB	ca. 0.3	37°C	5 hr
14	76-168	TB or LB	ca. 0.3	37°C	3 hr or 16-18 hr
15	105-209	M9ZB	ca. 0.3	37°C	3 hr
16	169-209	no expression			
17	105-168	LB	ca. 0.3	37°C	3-5 hr
18	99-209	LB	ca. 0.3	37°C	3 hr
19	137-204	LB	ca. 0.3	37°C	3 hr
20	137-180	LB	ca. 0.3	37°C	16-18 hr.
21	105-180	LB	ca. 0.3	37°C	3 hr
22	150-209	no expression			
23	150-180	no expression			

- 5 General expression method: *Escherichia coli* BL21(DE3) strains containing the hrpN subclones were grown in Luria broth medium (5g/L Difco Yeast extract, 10 g/L Difco Tryptone, 5g/L NaCl, and 1 mM NaOH) containing 30 µg/ml of kanamycin at 37°C overnight. The bacteria were then inoculated into 100 volumes of the same medium and grown at 37°C to an OD₆₂₀ of 0.6-0.8. The bacteria were then inoculated into 250 volumes of the growth medium and grown at 37°C to a specific induction OD₆₂₀. One milli molar IPTG was then added and the cultures grown at an optimal temperature for protein expression, and harvested at a particular time for recovery of the highest level of protein.
- 10

Example 18 - Small scale purification of harpin truncated proteins (verification of expression)

15

- A 50 ml culture of a hrpN clone was grown as above to induce expression of the truncated protein. Upon harvesting of the culture, 1.5 ml of the cell suspension were centrifuged at 14,000 rpm for 5 minutes, re-suspended in urea lysis buffer (8 M urea, 0.1 M Na₂HPO₄, and 0.01 M Tris -- pH 8.0), incubated at room temperature for 10 minutes, then centrifuged again at 14,000 rpm for 10 minutes, and the supernatant
- 20

- 49 -

saved. A 50 µl aliquot of a 50% slurry of an equilibrated (His)₆-binding nickel agarose resin was added to the supernatant and mixed at 4°C for one hour. The nickel agarose was then washed three times with urea washing buffer (8 M urea, 0.1 M Na₂HPO₄, and 0.01 M Tris -- pH 6.3), centrifuging at 5,000 rpm for five minutes
5 between washings. The protein was eluted from the resin with 50 µl of urea elution buffer (8 M urea, 0.1 M Na₂HPO₄, 0.01 M Tris, and 0.1 M EDTA -- pH 6.3). The eluate was run on a 4-20%, a 16%, or a 10-20% Tris-Glycine pre-cast gel depending upon the size of the truncated protein to verify the expression.

Example 19 - Induction of HR in tobacco

10

A 1.5 ml aliquot from the 50 ml cultures grown for small scale purification of the truncated proteins was centrifuged at 14,000 rpm for four minutes and re-suspended in an equal volume of 5 mM potassium phosphate buffer, pH 6.8. The cell suspension was sonicated for ca. 30 seconds then diluted 1:2 and 1:10 with
15 phosphate buffer. Both dilutions plus the neat cell lysate were infiltrated into the fourth to ninth leaves of 10-15 leaf tobacco plants by making a hole in single leaf panes and infiltrating the bacterial lysate into the intercellular leaf space using a syringe without a needle. The HR response was recorded 24-48 hr post infiltration. Tobacco (*Nicotiana tabacum* v. Xanthi) seedlings were grown in an environmental
20 chamber at 20-25°C with a photoperiod of 12-h light /12-h dark and ca. 40% RH. Cell lysate was used for the initial HR assays (in order to screen the truncated proteins for HR activity) as the small scale urea purification yielded very little protein which was denatured due to the purification process.

Example 20 - Large scale native purification of harpin truncated proteins for comprehensive biological activity assays

25

Six 500 ml cultures of a hrpN clone were grown as described earlier to induce expression of the truncated protein. Upon harvesting of the culture the cells were
30 centrifuged at 7,000 rpm for 5 minutes, re-suspended in imidazole lysis buffer (5 mM imidazole, 0.5 M NaCl, 20 mM Tris) plus Triton X-100 at 0.05% and lysozyme at 0.1 mg/ml, and incubated at 30°C for 15 minutes, sonicated for two minutes, then centrifuged again at 15,000 rpm for 20 minutes, and the supernatant was saved. A 4

- 50 -

ml aliquot of a 50% slurry of an equilibrated (His)₆-binding nickel agarose resin was added to the supernatant and mixed at 4°C for ca. four hours. The nickel agarose was then washed three times with imidazole washing buffer (20 mM imidazole, 0.5 M NaCl, and 20 mM Tris), centrifuging at 5,000 rpm for five minutes between washings, then placed in a disposable chromatography column. The column was centrifuged at 1100 rpm for one minute to remove any residual wash buffer and then the protein was eluted from the resin with 4 ml of imidazole elution buffer (1 M imidazole, 0.5 M NaCl, and 20 mM Tris) by incubating the column with the elution buffer for ten minutes at room temperature and then centrifuging the column at 1100 rpm for one minute. The eluate was run on a 4-20%, a 16%, or a 10-20% Tris-Glycine pre-cast gel depending upon the size of the truncated protein to verify the expression. The concentration of the proteins was determined by comparison of the protein bands with a standard protein in the Mark 12 molecular weight marker.

Example 21 - Large scale urea purification of harpin truncated proteins for comprehensive biological activity assay

The procedure was the same as the large scale native purification except that urea lysis buffer, washing buffer, and elution buffer were used, and the cells were not sonicated as in the native purification. After purification, the protein was renatured by dialyzing against lower and lower concentrations of urea over an eight hour period, then dialyzing overnight against 10 mM Tris/20 mM NaCl. The renaturing process caused the N-terminal proteins to precipitate. The precipitated 1-168 protein was solubilized by the addition of 100 mM Tris-HCl at pH 10.4 then heating the protein at 30°C for ca. one hour. The concentration of the protein was determined by comparison of the protein bands with a standard protein in the Mark 12 molecular weight marker. The 1-75 and 1-104 protein fragments were not successfully solubilized using this strategy so they were sonicated in 100 mM Tris-HCl at pH 10.4 to solubilize as much of the protein as possible and expose the active sites of the protein for the biological activity assays.

- 51 -

Example 22 - Expression of harpin truncated proteins

The small scale expression and purification of the fragment proteins was done to screen for expression and HR activity (Table 3).

5

Table 3

Expression and HR activity of harpin truncated proteins (small scale screening)

Fragment #	Amino Acids (SEQ. ID. No. 23)	Expression	HR activity
1(+control)	1-403	+	+
2(- control)	-	background protein only	-
3	105-403	+	+
4	169-403	+	-
5	210-403	+	-
6	267-403	+	-
7	343-403	+/-	-
8	1-75	+	-
9	1-104	+	+/-
10	1-168	+	+
11	1-266	+	+
12	1-342	+	+
13	76-209	+	+
14	76-168	+	-
15	105-209	+	+
16	169-209	-	-
17	105-168	+	-
18	99-209	+	+
19	137-204	+	+
20	137-180	+	+
21	105-180	+	+
22	150-209	-	-
23	150-180	-	-

- 10 All of the cloned fragment proteins were expressed to a certain degree except for three small fragments (amino acids 169-209, 150-209, and 150-180). The fragments were expressed at varying levels. Fragments 210-403 and 267-403 were expressed very well, yielding a high concentration of protein from a small scale purification, resulting in a substantial protein band on SDS gel electrophoresis. Other fragments (such as
- 15 a.a. 1-168 and 1-104) produced much less protein, resulting in faint protein bands upon electrophoresis. It was difficult to determine whether fragment 343-403, the smallest C-terminal protein, was expressed, as there were several background proteins apparent on the gel, in addition to the suspected 343-403 protein. The positive and

- 52 -

negative control proteins, consisting of the full length harpin protein and only background proteins, respectively, were tested for expression and HR activity as well.

The large scale expression and purification of the fragment proteins was done to determine the level of expression and titer of the HR activity (Table 4).

5

Table 4

Expression level and HR titer of harpin truncated proteins (large sale purification)

Fragment #	Amino acids (SEQ. ID. No. 23)	Expression	HR titer
1 (+ control)	1-403	3.7 mg/ml	5-7 µg/ml
2 (- control)	-	-	1:2 dilution
4	169-403	2 mg/ml	-
5	210-403	5 mg/ml	-
6	267-403	4 mg/ml	-
7	343-402	200µg/ml	-
8	1-75	50µg/ml	-
9	1-104	50µg/ml	3 µg/ml (1:16 dilution)
10	1-168	1 mg/ml	1 µg/ml
13	76-209	2.5 mg/ml	5 µg/ml
14	76-168	2 mg/ml	-
15	105-209	5 mg/ml	5-10µg/ml
17	105-168	250µg/ml	-
19	137-204	3.6 mg/ml	3.5 µg/ml
20	137-180	250 µg/ml	16 µg/ml

10

Not all of the proteins were expressed in large scale due to time constraints. The truncated proteins deemed to be the most important in characterizing harpin were chosen. The positive control (full length harpin) was expressed in a relatively high level at 3.7 mg/ml. All of the C-terminal proteins were expressed at relatively high levels from 2-5 mg/ml, except for fragment 343-403 as discussed earlier. The N-terminal fragments were expressed very well also, however, during the purification process, the protein precipitated and very little was resolubilized. The concentrations in Table 3 reflect only the solubilized protein. The internal fragments were expressed in the range of 2-3.6 mg/ml. It was extremely difficult to determine the concentration of fragment 105-168 (it was suspected that the concentration was much higher than indicated), as the protein bands on the SDS gel were large, but poorly stained. The negative control contained several background proteins as expected, but no obviously induced dominant protein.

15

20

- 53 -

Example 23 - Induction of HR in tobacco

The full length positive control protein elicited HR down to only 5-7µg/ml. The negative control (pET 28) imidazole purified "protein" - which contained only background proteins - elicited an HR response down to the 1:2 dilution, which lowered the sensitivity of the assay as the 1:1 and 1:2 dilutions could not be used. This false HR was likely due an affinity of the imidazole used in the purification process to bind to one or several of the background proteins, thereby not completely dialyzing out. Imidazole at a concentration of ca. 60 mM did elicit a false HR response.

One definitive domain encompassed a small internal region of the protein from a.a. 137-180 (SEQ. ID. No. 23), a mere 44 a.a, is identified as the smallest HR domain. The other potential HR domain is thought to be located in the N-terminus of the protein from a.a. 1-104 (possibly a.a. 1-75) (SEQ. ID. No. 23). It was difficult to confirm or narrow down the N-terminus HR domain due to the difficulties encountered in purifying these fragment proteins. The N-terminus fragment proteins had to be purified with urea as no protein was recovered when the native purification process was used. Consequently, these proteins precipitated during the renaturing process and were difficult or nearly impossible to get back into solution, thereby making it hard to run the proteins through the HR assay, as only soluble protein is able to elicit HR. Difficulty narrowing the N-terminus HR domain was only compounded by the fact that the negative control elicited false HR at the low dilution levels thereby reducing the sensitivity of the assay.

The internal domain proteins elicited an HR response between 5 and 10 µg/ml of protein like the positive control, and the N-terminus domain proteins elicited an HR response between 1 and 3 µg/ml, lower than the positive control.

Surprisingly, when the internal HR domain was cleaved between a.a. 168 and 169 (fragments 76-168 and 105-168) (SEQ. ID. No. 23) the fragment lost its HR activity. This suggests that the HR activity of fragment 1-168 (SEQ. ID. No. 23) should not be attributed to the internal HR domain, but rather to some other domain, leading to the assumption that there was likely a second HR domain to be found in the N-terminal region of the protein. However, as discussed earlier it was difficult to confirm this assumption.

- 54 -

The harpin C-terminus (a.a. 210-403 (SEQ. ID. No. 23)) did not contain an HR domain. It did not elicit HR at a detectable level using the current HR assay. Even the large C-terminal fragment from a.a. 169-403 (SEQ. ID. No. 23) did not elicit HR even though it contained part of the internal HR domain. As stated above, the
5 protein between a.a. 168 and 169 (SEQ. ID. No. 23) causes a loss of HR activity.

Because some of the small cloned proteins with 61 a.a. or less were not expressed, several oligopeptides were synthesized with 30 a.a. to narrow down the functional region of the internal HR domain. The oligopeptides were synthesized within the range of a.a. 121-179 (SEQ. ID. No. 23). However, these oligos did not
10 elicit the HR response. It was not expected that there would be an HR response from oligos 137-166, 121-150, and 137-156 (SEQ. ID. No. 23) as these fragments did not contain the imperative amino acids 168 and 169 (SEQ. ID. No. 23). It was expected that the oligo 150-179 (SEQ. ID. No. 23) would elicit an HR response. It is possible that 30 a.a. is too small for the protein to elicit any activity due to a lack of folding
15 and, therefore, a lack of binding or that during the synthesis of the peptides important amino acids were missed (either in the process, or simply by the choice of which 30 amino acids to synthesize) and, therefore, the fragments would not be able to elicit HR. It is also possible, although unlikely, that these small proteins would have undergone some form of post-translational modification within the *E.coli* cell that
20 they did not contain when synthesized and, therefore, were not able to elicit an HR response.

Example 24 – Biological Activity of HR Inducing Fragments

25 The two N-terminal harpin fragments spanning nucleotides 1-104 and nucleotides 1-168 of the nucleic acid of SEQ. ID. No. 24 were effective at inducing resistance of tobacco against TMV, in a similar manner as the full length harpin protein. The internal fragments spanning nucleotides 76-209 and nucleotides 105-209 of the nucleic acid of SEQ. ID. No. 24 were also effective at inducing TMV
30 resistance. In addition, these same four fragments conferred plant growth enhancement ("PGE") in tomato increasing the height of the plants from 4-19% taller than the buffer control plants. The full length harpin protein induced growth

- 55 -

enhancement of 6% greater than the buffer. The negative control did not induce TMV resistance or growth enhancement.

Table 5

5 TMV resistance and PGE activity of HR inducing fragments derived from harpin

Fragment #	Amino acids (SEQ. ID. No. 23)	HR activity	TMV resistance	PGE ht > buffer
<i>1 (+ control)</i>	1-403	+	+	6%
<i>2 (- control)</i>	-	-	-	-2%
<i>9</i>	1-104	+	+	4-8%
<i>10</i>	1-168	+	+	5-13%
<i>13</i>	76-209	+	+	4-18%
<i>15</i>	105-209	+	+	6-19%

Although the invention has been described in detail for the purpose of illustration, it is understood that such detail is solely for that purpose, and variations
10 can be made therein by those skilled in the art without departing from the spirit and scope of the invention which is defined by the following claims.

- 56 -

SEQUENCE LISTING

(1) GENERAL INFORMATION:

- (i) APPLICANT: Cornell Research Foundation, Inc. and EDEN Bioscience Corporation
- (ii) TITLE OF INVENTION: HYPERSENSITIVE RESPONSE ELICITOR
 FRAGMENTS ELICITING A HYPERSENSITIVE
 RESPONSE AND USES THEREOF
- (iii) NUMBER OF SEQUENCES: 30
- (iv) CORRESPONDENCE ADDRESS:
 - (A) ADDRESSEE: Nixon, Hargrave, Devans & Doyle LLP
 - (B) STREET: Clinton Square, P.O. Box 1051
 - (C) CITY: Rochester
 - (D) STATE: New York
 - (E) COUNTRY: U.S.A.
 - (F) ZIP: 14603
- (v) COMPUTER READABLE FORM:
 - (A) MEDIUM TYPE: Floppy disk
 - (B) COMPUTER: IBM PC compatible
 - (C) OPERATING SYSTEM: PC-DOS/MS-DOS
 - (D) SOFTWARE: PatentIn Release #1.0, Version #1.30
- (vi) CURRENT APPLICATION DATA:
 - (A) APPLICATION NUMBER:
 - (B) FILING DATE:
 - (C) CLASSIFICATION:
- (vii) PRIOR APPLICATION DATA:
 - (A) APPLICATION NUMBER: US 60/048,109
 - (B) FILING DATE: 30-MAY-1997
- (viii) ATTORNEY/AGENT INFORMATION:
 - (A) NAME: Goldman, Michael L.
 - (B) REGISTRATION NUMBER: 30,727
 - (C) REFERENCE/DOCKET NUMBER: 19603/1302
- (ix) TELECOMMUNICATION INFORMATION:
 - (A) TELEPHONE: (716) 263-1304
 - (B) TELEFAX: (716) 263-1600

(2) INFORMATION FOR SEQ ID NO:1:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 31 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

- 57 -

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

GGGAATTCAT ATGAGTCTGA ATACAAGTGG G

31

(2) INFORMATION FOR SEQ ID NO:2:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 31 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

GGGAATTCAT ATGGGCGGTG GCTTAGGCGG T

31

(2) INFORMATION FOR SEQ ID NO:3:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 29 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

GGCATATGTC GAACGCGCTG AACGATATG

29

(2) INFORMATION FOR SEQ ID NO:4:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 31 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

GGGAATTCAT ATGTTAGGCG GTTCGCTGAA C

31

(2) INFORMATION FOR SEQ ID NO:5:

- 58 -

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 29 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

- (ii) MOLECULE TYPE: cDNA

- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

GGCATATGCT GAACACGCTG GGCTCGAAA

29

- (2) INFORMATION FOR SEQ ID NO:6:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 29 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

- (ii) MOLECULE TYPE: cDNA

- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

GGCATATGTC AACGTCCCAA AACGACGAT

29

- (2) INFORMATION FOR SEQ ID NO:7:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 27 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

- (ii) MOLECULE TYPE: cDNA

- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:

GGCATATGTC CACCTCAGAC TCCAGCG

27

- (2) INFORMATION FOR SEQ ID NO:8:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 34 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

- 59 -

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:8:

GGGAATTCAT ATGCAAAGCC TGTTTGGTGA TGGG

34

(2) INFORMATION FOR SEQ ID NO:9:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 31 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:9:

GGGAATTCAT ATGGGTAATG GTCTGAGCAA G

31

(2) INFORMATION FOR SEQ ID NO:10:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 31 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:10:

GGGAATTCAT ATGAAAGCGG GCATTCAGGC G

31

(2) INFORMATION FOR SEQ ID NO:11:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 34 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:11:

GGGAATTCAT ATGACACCAG CCAGTATGGA GCAG

34

- 60 -

(2) INFORMATION FOR SEQ ID NO:12:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 31 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:12:

GCAAGCTTAA CAGCCCACCA CCGCCCATCA T

31

(2) INFORMATION FOR SEQ ID NO:13:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 31 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:13:

GCAAGCTTAA ATCGTTCAGC GCGTTCGACA G

31

(2) INFORMATION FOR SEQ ID NO:14:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 34 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:14:

GCAAGCTTAA TATCTCGCTG AACATCTTCA GCAG

34

(2) INFORMATION FOR SEQ ID NO:15:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 30 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

- 61 -

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:15:

GCAAGCTTAA GGTGCCATCT TGCCCATCAC

30

(2) INFORMATION FOR SEQ ID NO:16:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 34 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:16:

GCAAGCTTAA ATCAGTGACT CCTTTTTTAT AGGC

34

(2) INFORMATION FOR SEQ ID NO:17:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 31 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:17:

GCAAGCTTAA CAGGCCCGAC AGCGCATCAG T

31

(2) INFORMATION FOR SEQ ID NO:18:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 31 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:18:

GCAAGCTTAA ACCGATACCG GTACCCACGG C

31

- 62 -

(2) INFORMATION FOR SEQ ID NO:19:

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 34 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:19:

GCAAGCTTAA TCCGTCGTCA TCTGGCTTGC TCAG

34

(2) INFORMATION FOR SEQ ID NO:20:

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 25 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:20:

GCAAGCTTAA GCCGCGCCCA GCTTG

25

(2) INFORMATION FOR SEQ ID NO:21:

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 338 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS:
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:21:

Met	Gln	Ile	Thr	Ile	Lys	Ala	His	Ile	Gly	Gly	Asp	Leu	Gly	Val	Ser
1					5				10					15	
Gly	Leu	Gly	Ala	Gln	Gly	Leu	Lys	Gly	Leu	Asn	Ser	Ala	Ala	Ser	Ser
				20				25						30	
Leu	Gly	Ser	Ser	Val	Asp	Lys	Leu	Ser	Ser	Thr	Ile	Asp	Lys	Leu	Thr
				35				40						45	

- 63 -

Ser Ala Leu Thr Ser Met Met Phe Gly Gly Ala Leu Ala Gln Gly Leu
 50 55 60
 Gly Ala Ser Ser Lys Gly Leu Gly Met Ser Asn Gln Leu Gly Gln Ser
 65 70 75 80
 Phe Gly Asn Gly Ala Gln Gly Ala Ser Asn Leu Leu Ser Val Pro Lys
 85 90 95
 Ser Gly Gly Asp Ala Leu Ser Lys Met Phe Asp Lys Ala Leu Asp Asp
 100 105 110
 Leu Leu Gly His Asp Thr Val Thr Lys Leu Thr Asn Gln Ser Asn Gln
 115 120 125
 Leu Ala Asn Ser Met Leu Asn Ala Ser Gln Met Thr Gln Gly Asn Met
 130 135 140
 Asn Ala Phe Gly Ser Gly Val Asn Asn Ala Leu Ser Ser Ile Leu Gly
 145 150 155 160
 Asn Gly Leu Gly Gln Ser Met Ser Gly Phe Ser Gln Pro Ser Leu Gly
 165 170 175
 Ala Gly Gly Leu Gln Gly Leu Ser Gly Ala Gly Ala Phe Asn Gln Leu
 180 185 190
 Gly Asn Ala Ile Gly Met Gly Val Gly Gln Asn Ala Ala Leu Ser Ala
 195 200 205
 Leu Ser Asn Val Ser Thr His Val Asp Gly Asn Asn Arg His Phe Val
 210 215 220
 Asp Lys Glu Asp Arg Gly Met Ala Lys Glu Ile Gly Gln Phe Met Asp
 225 230 235 240
 Gln Tyr Pro Glu Ile Phe Gly Lys Pro Glu Tyr Gln Lys Asp Gly Trp
 245 250 255
 Ser Ser Pro Lys Thr Asp Asp Lys Ser Trp Ala Lys Ala Leu Ser Lys
 260 265 270
 Pro Asp Asp Asp Gly Met Thr Gly Ala Ser Met Asp Lys Phe Arg Gln
 275 280 285
 Ala Met Gly Met Ile Lys Ser Ala Val Ala Gly Asp Thr Gly Asn Thr
 290 295 300
 Asn Leu Asn Leu Arg Gly Ala Gly Gly Ala Ser Leu Gly Ile Asp Ala
 305 310 315 320
 Ala Val Val Gly Asp Lys Ile Ala Asn Met Ser Leu Gly Lys Leu Ala
 325 330 335
 Asn Ala

- 64 -

(2) INFORMATION FOR SEQ ID NO:22:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 2141 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:22:

CGATTTTACC CGGGTGAACG TGCTATGACC GACAGCATCA CGGTATTCTGA CACCGTTACG	60
GCGTTTATGG CCGCGATGAA CCGGCATCAG GCGGCGCGCT GGTGCGCCGA ATCCGGCGTC	120
GATCTGGTAT TTCAGTTTGG GGACACCGGG CGTGAACCTCA TGATGCAGAT TCAGCCGGGG	180
CAGCAATATC CCGGCATGTT GCGCAGCTG CTCGCTCGTC GTTATCAGCA GCGGCAGAG	240
TGCGATGGCT GCCATCTGTG CCTGAACGGC AGCGATGTAT TGATCCTCTG GTGGCCGCTG	300
CCGTCGGATC CCGGCAGTTA TCCGCAGGTG ATCGAACGTT TGTTTGAACT GCGGGGAATG	360
ACGTTGCCGT CGCTATCCAT AGCACCGACG GCGCGTCCGC AGACAGGGAA CCGACGCGCC	420
CGATCATTA GATAAAGGCG GCTTTTTTTT TTGCAAAACG GTAACGGTGA GGAACCGTTT	480
CACCGTCGGC GTCATCTAGT AACAAGTATC CATCATGATG CCTACATCGG GATCGGCGTG	540
GGCATCCGTT GCAGATACTT TTGCGAACAC CTGACATGAA TGAGGAAACG AAATTATGCA	600
AATTACGATC AAAGCGCACA TCGGCGGTGA TTTGGGCGTC TCCGGTCTGG GGCTGGGTGC	660
TCAGGGACTG AAAGGACTGA ATTCCGCGGC TTCATCGCTG GGTTCAGCG TGGATAAACT	720
GAGCAGCACC ATCGATAAGT TGACCTCCGC GCTGACTTCG ATGATGTTTG GCGGCGCGCT	780
GGCGCAGGGG CTGGGCGCCA GCTCGAAGGG GCTGGGGATG AGCAATCAAC TGGGCCAGTC	840
TTTCGGCAAT GCGCGCGAGG GTGCGAGCAA CCTGCTATCC GTACCGAAAT CCGGCGGCGA	900
TGCGTTGTCA AAAATGTTTG ATAAAGCGCT GGACGATCTG CTGGGTCATG ACACCGTGAC	960
CAAGCTGACT AACCAGAGCA ACCAACTGGC TAATTCAATG CTGAACGCCA GCCAGATGAC	1020
CCAGGGTAAT ATGAATGCGT TCGGCAGCGG TGTGAACAAC GCACTGTCGT CCATTCTCGG	1080
CAACGGTCTC GGCCAGTCGA TGAGTGCTT CTCTCAGCCT TCTCTGGGGG CAGGCGGCTT	1140
GCAGGGCCTG AGCGGCGCGG GTGCATTCAA CCAGTTGGGT AATGCCATCG GCATGGGCGT	1200
GGGGCAGAAT GCTGCGCTGA GTGCGTTGAG TAACGTCAGC ACCCACGTAG ACGGTAACAA	1260

- 65 -

```

CCGCCACTTT GTAGATAAAG AAGATCGCGG CATGGCGAAA GAGATCGGCC AGTTTATGGA      1320
TCAGTATCCG GAAATATTTC GTAAACCGGA ATACCAGAAA GATGGCTGGA GTTCGCCGAA      1380
GACGGACGAC AAATCCTGGG CTAAAGCGCT GAGTAAACCG GATGATGACG GTATGACCGG      1440
CGCCAGCATG GACAAATTCC GTCAGGCGAT GGGTATGATC AAAAGCGCGG TGGCGGGTGA      1500
TACCGGCAAT ACCAACCTGA ACCTGCGTGG CGCGGGCGGT GCATCGCTGG GTATCGATGC      1560
GGCTGTCGTC GGCGATAAAA TAGCCAACAT GTCGCTGGGT AAGCTGGCCA ACGCCTGATA      1620
ATCTGTGCTG GCCTGATAAA GCGGAAACGA AAAAAGAGAC GGGGAAGCCT GTCTCTTTTC      1680
TTATTATGCG GTTTATGCGG TTACCTGGAC CGGTTAATCA TCGTCATCGA TCTGGTACAA      1740
ACGCACATTT TCCCGTTCAT TCGCGTCGTT ACGCGCCACA ATCGCGATGG CATCTTCCTC      1800
GTCGCTCAGA TTGCGCGGCT GATGGGGAAC GCCGGGTGGA ATATAGAGAA ACTCGCCGGC      1860
CAGATGGAGA CACGTCTGCG ATAAATCTGT GCCGTAACGT GTTCTATATCC GCCCCTTTAG      1920
CAGATAGATT GCGGTTTCGT AATCAACATG GTAATGCGGT TCCGCCTGTG CGCCGGCCGG      1980
GATCACCACA ATATTCATAG AAAGCTGTCT TGCACCTACC GTATCGCGGG AGATACCGAC      2040
AAAATAGGGC AGTTTTTTCG TGGTATCCGT GGGGTGTTCC GGCCTGACAA TCTTGAGTTG      2100
GTTCGTCATC ATCTTTCTCC ATCTGGGCGA CCTGATCGGT T                                2141

```

(2) INFORMATION FOR SEQ ID NO:23:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 403 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS:
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:23:

```

Met Ser Leu Asn Thr Ser Gly Leu Gly Ala Ser Thr Met Gln Ile Ser
1              5              10              15

Ile Gly Gly Ala Gly Gly Asn Asn Gly Leu Leu Gly Thr Ser Arg Gln
                20              25              30

Asn Ala Gly Leu Gly Gly Asn Ser Ala Leu Gly Leu Gly Gly Gly Asn
35              40              45

```

- 66 -

Gln Asn Asp Thr Val Asn Gln Leu Ala Gly Leu Leu Thr Gly Met Met
 50 55 60

Met Met Met Ser Met Met Gly Gly Gly Gly Leu Met Gly Gly Gly Leu
 65 70 75 80

Gly Gly Gly Leu Gly Asn Gly Leu Gly Gly Ser Gly Gly Leu Gly Glu
 85 90 95

Gly Leu Ser Asn Ala Leu Asn Asp Met Leu Gly Gly Ser Leu Asn Thr
 100 105 110

Leu Gly Ser Lys Gly Gly Asn Asn Thr Thr Ser Thr Thr Asn Ser Pro
 115 120 125

Leu Asp Gln Ala Leu Gly Ile Asn Ser Thr Ser Gln Asn Asp Asp Ser
 130 135 140

Thr Ser Gly Thr Asp Ser Thr Ser Asp Ser Ser Asp Pro Met Gln Gln
 145 150 155 160

Leu Leu Lys Met Phe Ser Glu Ile Met Gln Ser Leu Phe Gly Asp Gly
 165 170 175

Gln Asp Gly Thr Gln Gly Ser Ser Ser Gly Gly Lys Gln Pro Thr Glu
 180 185 190

Gly Glu Gln Asn Ala Tyr Lys Lys Gly Val Thr Asp Ala Leu Ser Gly
 195 200 205

Leu Met Gly Asn Gly Leu Ser Gln Leu Leu Gly Asn Gly Gly Leu Gly
 210 215 220

Gly Gly Gln Gly Gly Asn Ala Gly Thr Gly Leu Asp Gly Ser Ser Leu
 225 230 235 240

Gly Gly Lys Gly Leu Gln Asn Leu Ser Gly Pro Val Asp Tyr Gln Gln
 245 250 255

Leu Gly Asn Ala Val Gly Thr Gly Ile Gly Met Lys Ala Gly Ile Gln
 260 265 270

Ala Leu Asn Asp Ile Gly Thr His Arg His Ser Ser Thr Arg Ser Phe
 275 280 285

Val Asn Lys Gly Asp Arg Ala Met Ala Lys Glu Ile Gly Gln Phe Met
 290 295 300

Asp Gln Tyr Pro Glu Val Phe Gly Lys Pro Gln Tyr Gln Lys Gly Pro
 305 310 315 320

Gly Gln Glu Val Lys Thr Asp Asp Lys Ser Trp Ala Lys Ala Leu Ser
 325 330 335

[illegible]

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1288 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:24:

AAGCTTCGGC	ATGGCACGTT	TGACCGTTGG	GTCGGCAGGG	TACGTTTGAA	TTATTTCATAA	60
GAGGAATACG	TTATGAGTCT	GAATACAAGT	GGGCTGGGAG	CGTCAACGAT	GCAAATTTCT	120
ATCGGCGGTG	CGGGCGGAAA	TAACGGGTTG	CTGGGTACCA	GTCGCCAGAA	TGCTGGGTTG	180
GGTGGCAATT	CTGCACTGGG	GCTGGGCGGC	GGTAATCAAA	ATGATACCGT	CAATCAGCTG	240
GCTGGCTTAC	TCACCGGCAT	GATGATGATG	ATGAGCATGA	TGGGCGGTGG	TGGGCTGATG	300
GGCGGTGGCT	TAGGCGGTGG	CTTAGGTAAT	GGCTTGGGTG	GCTCAGGTGG	CCTGGGCGAA	360
GGACTGTCTGA	ACGCGCTGAA	CGATATGTTA	GGCGGTTCGC	TGAACACGCT	GGGCTCGAAA	420
GGCGGCAACA	ATACCACTTC	AACAACAAAT	TCCCCGCTGG	ACCAGGCGCT	GGGTATTAAC	480
TCAACGTCCC	AAAACGACGA	TTCCACCTCC	GGCACAGATT	CCACCTCAGA	CTCCAGCGAC	540
CCGATGCAGC	AGCTGCTGAA	GATGTTTCAGC	GAGATAATGC	AAAGCCTGTT	TGGTGATGGG	600
CAAGATGGCA	CCCAGGGCAG	TTCCTCTGGG	GGCAAGCAGC	CGACCGAAGG	CGAGCAGAAC	660
GCCTATAAAA	AAGGAGTCAC	TGATGCGCTG	TCGGGCCTGA	TGGGTAATGG	TCTGAGCCAG	720
CTCCTTGGCA	ACGGGGGACT	GGGAGGTGGT	CAGGGCGGTA	ATGCTGGCAC	GGGTCTTGAC	780

- 68 -

```

GGTTCGTCGC TGGGCGGCAA AGGGCTGCAA AACCTGAGCG GGCCGGTGGA CTACCAGCAG      840
TTAGGTAACG CCGTGGGTAC CGGTATCGGT ATGAAAGCGG GCATTTCAGGC GCTGAATGAT      900
ATCGGTACGC ACAGGCACAG TTCAACCCGT TCTTTCGTCA ATAAAGGCGA TCGGGCGATG      960
GCGAAGGAAA TCGGTCAGTT CATGGACCAG TATCCTGAGG TGTTTGGCAA GCCGCAGTAC     1020
CAGAAAGGCC CGGGTCAGGA GGTGAAAACC GATGACAAAT CATGGGCAAA AGCACTGAGC     1080
AAGCCAGATG ACGACGGAAT GACACCAGCC AGTATGGAGC AGTTCAACAA AGCCAAGGGC     1140
ATGATCAAAA GGCCCATGGC GGGTGATACC GGCAACGGCA ACCTGCAGGC ACGCGGTGCC     1200
GGTGGTTCTT CGCTGGGTAT TGATGCCATG ATGGCCGGTG ATGCCATTAA CAATATGGCA     1260
CTTGGCAAGC TGGGCGCGGC TTAAGCTT                                     1288

```

(2) INFORMATION FOR SEQ ID NO:25:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 341 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS:
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:25:

```

Met Gln Ser Leu Ser Leu Asn Ser Ser Ser Leu Gln Thr Pro Ala Met
1           5           10           15
Ala Leu Val Leu Val Arg Pro Glu Ala Glu Thr Thr Gly Ser Thr Ser
20          25          30
Ser Lys Ala Leu Gln Glu Val Val Val Lys Leu Ala Glu Glu Leu Met
35          40          45
Arg Asn Gly Gln Leu Asp Asp Ser Ser Pro Leu Gly Lys Leu Leu Ala
50          55          60
Lys Ser Met Ala Ala Asp Gly Lys Ala Gly Gly Gly Ile Glu Asp Val
65          70          75          80
Ile Ala Ala Leu Asp Lys Leu Ile His Glu Lys Leu Gly Asp Asn Phe
85          90          95
Gly Ala Ser Ala Asp Ser Ala Ser Gly Thr Gly Gln Gln Asp Leu Met
100         105         110
Thr Gln Val Leu Asn Gly Leu Ala Lys Ser Met Leu Asp Asp Leu Leu
115         120         125

```

- 69 -

Thr Lys Gln Asp Gly Gly Thr Ser Phe Ser Glu Asp Asp Met Pro Met
 130 135 140
 Leu Asn Lys Ile Ala Gln Phe Met Asp Asp Asn Pro Ala Gln Phe Pro
 145 150 155 160
 Lys Pro Asp Ser Gly Ser Trp Val Asn Glu Leu Lys Glu Asp Asn Phe
 165 170 175
 Leu Asp Gly Asp Glu Thr Ala Ala Phe Arg Ser Ala Leu Asp Ile Ile
 180 185 190
 Gly Gln Gln Leu Gly Asn Gln Gln Ser Asp Ala Gly Ser Leu Ala Gly
 195 200 205
 Thr Gly Gly Gly Leu Gly Thr Pro Ser Ser Phe Ser Asn Asn Ser Ser
 210 215 220
 Val Met Gly Asp Pro Leu Ile Asp Ala Asn Thr Gly Pro Gly Asp Ser
 225 230 235 240
 Gly Asn Thr Arg Gly Glu Ala Gly Gln Leu Ile Gly Glu Leu Ile Asp
 245 250 255
 Arg Gly Leu Gln Ser Val Leu Ala Gly Gly Gly Leu Gly Thr Pro Val
 260 265 270
 Asn Thr Pro Gln Thr Gly Thr Ser Ala Asn Gly Gly Gln Ser Ala Gln
 275 280 285
 Asp Leu Asp Gln Leu Leu Gly Gly Leu Leu Leu Lys Gly Leu Glu Ala
 290 295 300
 Thr Leu Lys Asp Ala Gly Gln Thr Gly Thr Asp Val Gln Ser Ser Ala
 305 310 315 320
 Ala Gln Ile Ala Thr Leu Leu Val Ser Thr Leu Leu Gln Gly Thr Arg
 325 330 335
 Asn Gln Ala Ala Ala
 340

(2) INFORMATION FOR SEQ ID NO:26:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1026 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

- 70 -

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:26:

ATGCAGAGTC TCAGTCTTAA CAGCAGCTCG CTGCAAACCC CGGCAATGGC CCTTGTCCTG	60
GTACGTCCTG AAGCCGAGAC GACTGGCAGT ACGTCGAGCA AGGCGCTTCA GGAAGTTGTC	120
GTGAAGCTGG CCGAGGAACT GATGCGCAAT GGTCAACTCG ACGACAGCTC GCCATTGGGA	180
AAACTGTTGG CCAAGTCGAT GGCCGCAGAT GGCAAGGCGG GCGGCGGTAT TGAGGATGTC	240
ATCGCTGCGC TGGACAAGCT GATCCATGAA AAGCTCGGTG ACAACTTCGG CGCGTCTGCG	300
GACAGCGCCT CGGGTACCGG ACAGCAGGAC CTGATGACTC AGGTGCTCAA TGGCCTGGCC	360
AAGTCGATGC TCGATGATCT TCTGACCAAG CAGGATGGCG GGACAAGCTT CTCCGAAGAC	420
GATATGCCGA TGCTGAACAA GATCGCGCAG TTCATGGATG ACAATCCCGC ACAGTTTCCC	480
AAGCCGGACT CGGGCTCCTG GGTGAACGAA CTCAAGGAAG ACAACTTCCT TGATGGCGAC	540
GAAACGGCTG CGTTCCGTTC GGCACTCGAC ATCATTGGCC AGCAACTGGG TAATCAGCAG	600
AGTGACGCTG GCAGTCTGGC AGGGACGGGT GGAGGTCTGG GCACTCCGAG CAGTTTTTCC	660
AACAATCGT CCGTGATGGG TGATCCGCTG ATCGACGCCA ATACCGGTCC CGGTGACAGC	720
GGCAATACCC GTGGTGAAGC GGGGCAACTG ATCGGCGAGC TTATCGACCG TGGCCTGCAA	780
TCGGTATTGG CCGGTGGTGG ACTGGGCACA CCCGTAAACA CCCCAGAC CCGTACGTCG	840
GCGAATGGCG GACAGTCCGC TCAGGATCTT GATCAGTTGC TGGGCGGCTT GCTGCTCAAG	900
GGCCTGGAGG CAACGCTCAA GGATGCCGGG CAAACAGGCA CCGACGTGCA GTCGAGCGCT	960
GCGCAAATCG CCACCTTGCT GGTCAGTACG CTGCTGCAAG GCACCCGCAA TCAGGCTGCA	1020
GCCTGA	1026

(2) INFORMATION FOR SEQ ID NO:27:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 344 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS:
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:27:

Met	Ser	Val	Gly	Asn	Ile	Gln	Ser	Pro	Ser	Asn	Leu	Pro	Gly	Leu	Gln
1					5				10				15		

- 71 -

Asn Leu Asn Leu Asn Thr Asn Thr Asn Ser Gln Gln Ser Gly Gln Ser
 20 25 30
 Val Gln Asp Leu Ile Lys Gln Val Glu Lys Asp Ile Leu Asn Ile Ile
 35 40 45
 Ala Ala Leu Val Gln Lys Ala Ala Gln Ser Ala Gly Gly Asn Thr Gly
 50 55 60
 Asn Thr Gly Asn Ala Pro Ala Lys Asp Gly Asn Ala Asn Ala Gly Ala
 65 70 75 80
 Asn Asp Pro Ser Lys Asn Asp Pro Ser Lys Ser Gln Ala Pro Gln Ser
 85 90 95
 Ala Asn Lys Thr Gly Asn Val Asp Asp Ala Asn Asn Gln Asp Pro Met
 100 105 110
 Gln Ala Leu Met Gln Leu Leu Glu Asp Leu Val Lys Leu Leu Lys Ala
 115 120 125
 Ala Leu His Met Gln Gln Pro Gly Gly Asn Asp Lys Gly Asn Gly Val
 130 135 140
 Gly Gly Ala Asn Gly Ala Lys Gly Ala Gly Gly Gln Gly Gly Leu Ala
 145 150 155 160
 Glu Ala Leu Gln Glu Ile Glu Gln Ile Leu Ala Gln Leu Gly Gly Gly
 165 170 175
 Gly Ala Gly Ala Gly Gly Ala Gly Gly Gly Val Gly Gly Ala Gly Gly
 180 185 190
 Ala Asp Gly Gly Ser Gly Ala Gly Gly Ala Gly Gly Ala Asn Gly Ala
 195 200 205
 Asp Gly Gly Asn Gly Val Asn Gly Asn Gln Ala Asn Gly Pro Gln Asn
 210 215 220
 Ala Gly Asp Val Asn Gly Ala Asn Gly Ala Asp Asp Gly Ser Glu Asp
 225 230 235 240
 Gln Gly Gly Leu Thr Gly Val Leu Gln Lys Leu Met Lys Ile Leu Asn
 245 250 255
 Ala Leu Val Gln Met Met Gln Gln Gly Gly Leu Gly Gly Gly Asn Gln
 260 265 270
 Ala Gln Gly Gly Ser Lys Gly Ala Gly Asn Ala Ser Pro Ala Ser Gly
 275 280 285
 Ala Asn Pro Gly Ala Asn Gln Pro Gly Ser Ala Asp Asp Gln Ser Ser
 290 295 300

- 72 -

Gly Gln Asn Asn Leu Gln Ser Gln Ile Met Asp Val Val Lys Glu Val
 305 310 315 320

Val Gln Ile Leu Gln Gln Met Leu Ala Ala Gln Asn Gly Gly Ser Gln
 325 330 335

Gln Ser Thr Ser Thr Gln Pro Met
 340

(2) INFORMATION FOR SEQ ID NO:28:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1035 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:28:

ATGTCAGTCG GAAACATCCA GAGCCCGTCG AACCTCCCGG GTCTGCAGAA CCTGAACCTC	60
AACACCAACA CCAACAGCCA GCAATCGGGC CAGTCCGTGC AAGACCTGAT CAAGCAGGTC	120
GAGAAGGACA TCCTCAACAT CATCGCAGCC CTCGTGCAGA AGGCCGCACA GTCGGCGGGC	180
GGCAACACCG GTAACACCGG CAACGCGCCG GCGAAGGACG GCAATGCCAA CGCGGGCGCC	240
AACGACCCGA GCAAGAACGA CCCGAGCAAG AGCCAGGCTC CGCAGTCGGC CAACAAGACC	300
GGCAACGTCG ACGACGCCAA CAACCAGGAT CCGATGCAAG CGCTGATGCA GCTGCTGGAA	360
GACCTGGTGA AGCTGCTGAA GGC GGCCCTG CACATGCAGC AGCCCGGCGG CAATGACAAG	420
GGCAACGGCG TGGGCGGTGC CAACGGCGCC AAGGGTGCCG GCGGCCAGGG CGGCCTGGCC	480
GAAGCGCTGC AGGAGATCGA GCAGATCCTC GCCCAGCTCG GCGGCGGCGG TGCTGGCGCC	540
GGCGGCGCGG GTGGCGGTGT CGGCGGTGCT GGTGGCGCGG ATGGCGGCTC CGGTGCGGGT	600
GGCGCAGGCG GTGCGAACGG CGCCGACGGC GGCAATGGCG TGAACGGCAA CCAGGCGAAC	660
GGCCCGCAGA ACGCAGGCGA TGTCAACGGT GCCAACGGCG CGGATGACGG CAGCGAAGAC	720
CAGGGCGGCC TCACCGGCGT GCTGCAAAAG CTGATGAAGA TCCTGAACGC GCTGGTGCAG	780
ATGATGCAGC AAGGCGGCCT CGGCGGCGGC AACCAGGCGC AGGGCGGCTC GAAGGGTGCC	840
GGCAACGCCT CGCCGGCTTC CGGCGCGAAC CCGGGCGCGA ACCAGCCCGG TTCGGCGGAT	900
GATCAATCGT CCGGCCAGAA CAATCTGCAA TCCAGATCA TGGATGTGGT GAAGGAGGTC	960

- 73 -

GTCCAGATCC TGCAGCAGAT GCTGGCGGCG CAGAACGGCG GCAGCCAGCA GTCCACCTCG 1020

ACGCAGCCGA TGTA 1035

(2) INFORMATION FOR SEQ ID NO:29:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 26 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS:
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:29:

Thr Leu Ile Glu Leu Met Ile Val Val Ala Ile Ile Ala Ile Leu Ala
1 5 10 15

Ala Ile Ala Leu Pro Ala Tyr Gln Asp Tyr
20 25

(2) INFORMATION FOR SEQ ID NO:30:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 20 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS:
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:30:

Ser Ser Gln Gln Ser Pro Ser Ala Gly Ser Glu Gln Gln Leu Asp Gln
 1 5 10 15
 Leu Leu Ala Met
 20

- 74 -

WHAT IS CLAIMED:

1. An isolated fragment of an *Erwinia* hypersensitive response elicitor protein or polypeptide, wherein said fragment elicits a hypersensitive response in
5 plants.
2. An isolated fragment according to claim 1, wherein the hypersensitive response elicitor protein or polypeptide is derived from an *Erwinia amylovora*, *Erwinia carotovora*, *Erwinia chrysanthemi*, or *Erwinia stewartii*.
10
3. An isolated fragment according to claim 2, wherein the hypersensitive response elicitor protein or polypeptide is derived from *Erwinia amylovora*.
4. An isolated fragment according to claim 3, wherein the fragment is
15 selected from the group consisting of a C-terminal fragment of the amino acid sequence of SEQ. ID. No. 23, an N-terminal fragment of the amino acid sequence of SEQ. ID. No. 23, and an internal fragment of the amino acid sequence of SEQ. ID. No. 23.
- 20 5. An isolated fragment according to claim 4, wherein the fragment is a C-terminal fragment of the amino acid sequence of SEQ. ID. No. 23 spanning amino acids 105 and 403 of SEQ. ID. No. 23.
6. An isolated fragment according to claim 4, wherein the fragment is an
25 N-terminal fragment of the amino acid sequence of SEQ. ID. No. 23 spanning the following amino acids of SEQ. ID. No. 23: 1 and 98, 1 and 104, 1 and 122, 1 and 168, 1 and 218, 1 and 266, 1 and 342, 1 and 321, and 1 and 372.
7. An isolated fragment according to claim 4, wherein the fragment is an
30 internal fragment of the amino acid sequence of SEQ. ID. No. 23 spanning the following amino acids of SEQ. ID. No. 23: 76 and 209, 105 and 209, 99 and 209, 137 and 204, 137 and 200, 109 and 204, 109 and 200, 137 and 180, and 105 and 180.

- 75 -

8. An isolated DNA molecule encoding a fragment according to claim 1.
9. An isolated DNA molecule according to claim 8, wherein the hypersensitive response elicitor protein or polypeptide is derived from an *Erwinia amylovora*, *Erwinia carotovora*, *Erwinia chrysanthemi*, or *Erwinia stewartii*.
10. An isolated DNA molecule according to claim 9, wherein the hypersensitive response elicitor protein or polypeptide is derived from *Erwinia amylovora*.
11. An isolated isolated DNA molecule according to claim 10, wherein the fragment is selected from the group consisting of a C-terminal fragment of the amino acid sequence of SEQ. ID. No. 23, an N-terminal fragment of the amino acid sequence of SEQ. ID. No. 23, and an internal fragment of the amino acid sequence of SEQ. ID. No. 23.
12. An isolated DNA molecule according to claim 10, wherein the fragment is a C-terminal fragment of the amino acid sequence of SEQ. ID. No. 23 spanning amino acids 105 and 403 of SEQ. ID. No. 23.
13. An isolated DNA molecule according to claim 10, wherein the fragment is an N-terminal fragment of the amino acid sequence of SEQ. ID. No. 23 spanning the following amino acids of SEQ. ID. No. 23: 1 and 98, 1 and 104, 1 and 122, 1 and 168, 1 and 218, 1 and 266, 1 and 342, 1 and 321, and 1 and 372.
14. An isolated DNA molecule according to claim 10, wherein the fragment is an internal fragment of the amino acid sequence of SEQ. ID. No. 23 spanning the following amino acids of SEQ. ID. No. 23: 76 and 209, 105 and 209, 99 and 209, 137 and 204, 137 and 200, 109 and 204, 109 and 200, 137 and 180, and 105 and 180.

- 76 -

15. An expression system transformed with a DNA molecule according to claim 8.
16. An expression system according to claim 15, wherein said DNA molecule is in proper sense orientation and correct reading frame.
17. A host cell transformed with a DNA molecule according to claim 8.
18. A host cell according to claim 17, wherein the host cell is selected from the group consisting of a plant cell and a bacterial cell.
19. A host cell according to claim 17, wherein the DNA molecule is transformed with an expression system.
20. A transgenic plant transformed with the DNA molecule of claim 8.
21. A transgenic plant according to claim 20, wherein the plant is selected from the group consisting of alfalfa, rice, wheat, barley, rye, cotton, sunflower, peanut, corn, potato, sweet potato, bean, pea, chicory, lettuce, endive, cabbage, brussel sprout, beet, parsnip, turnip, cauliflower, broccoli, turnip, radish, spinach, onion, garlic, eggplant, pepper, celery, carrot, squash, pumpkin, zucchini, cucumber, apple, pear, melon, citrus, strawberry, grape, raspberry, pineapple, soybean, tobacco, tomato, sorghum, and sugarcane.
22. A transgenic plant according to claim 20, wherein the plant is selected from the group consisting of *Arabidopsis thaliana*, *Saintpaulia*, petunia, pelargonium, poinsettia, chrysanthemum, carnation, and zinnia.
23. A transgenic plant seed transformed with the DNA molecule of claim 8.

- 77 -

24. A transgenic plant seed according to claim 23, wherein the plant seed is selected from the group consisting of alfalfa, rice, wheat, barley, rye, cotton, sunflower, peanut, corn, potato, sweet potato, bean, pea, chicory, lettuce, endive, cabbage, brussel sprout, beet, parsnip, turnip, cauliflower, broccoli, turnip, radish, 5 spinach, onion, garlic, eggplant, pepper, celery, carrot, squash, pumpkin, zucchini, cucumber, apple, pear, melon, citrus, strawberry, grape, raspberry, pineapple, soybean, tobacco, tomato, sorghum, and sugarcane.

25. A transgenic plant seed according to claim 23, wherein the plant seed 10 is selected from the group consisting of *Arabidopsis thaliana*, *Saintpaulia*, petunia, pelargonium, poinsettia, chrysanthemum, carnation, and zinnia.

26. A method of imparting disease resistance to plants comprising:
applying a fragment of a hypersensitive response elicitor protein or 15 polypeptide, which fragment elicits a hypersensitive response, in a non-infectious form to a plant or plant seed under conditions impart disease resistance.

27. A method according to claim 26, wherein plants are treated during said 20 applying.

28. A method according to claim 26, wherein plant seeds are treated during said applying, said method further comprising:

planting the seeds treated with the fragment of the hypersensitive response elicitor in natural or artificial soil and 25 propagating plants from the seeds planted in the soil.

29. A method of enhancing plant growth comprising:
applying a fragment of a hypersensitive response elicitor protein or polypeptide, which fragment elicits a hypersensitive response, in a non-infectious 30 form to a plant or plant seed under conditions effective to enhance plant growth.

- 78 -

30. A method according to claim 29, wherein plants are treated during said applying.

31. A method according to claim 29, wherein plant seeds are treated during said applying, said method further comprising:

planting the seeds treated with the fragment of the hypersensitive response elicitor in natural or artificial soil and propagating plants from the seeds planted in the soil.

32. A method of insect control for plants comprising: applying a fragment of a hypersensitive response elicitor protein or polypeptide, which fragment elicits a hypersensitive response, in a non-infectious form to a plant or plant seed under conditions effective to control insects.

33. A method according to claim 32, wherein plants are treated during said applying.

34. A method according to claim 32, wherein plant seeds are treated during said applying, said method further comprising:

planting the seeds treated with the fragment of the hypersensitive response elicitor in natural or artificial soil and propagating plants from the seeds planted in the soil.

35. A method of imparting disease resistance to plants comprising: providing a transgenic plant or plant seed transformed with a DNA molecule which encodes a fragment of a hypersensitive response elicitor protein or polypeptide, which fragment elicits a hypersensitive response, and

growing the transgenic plant or transgenic plants produced from the transgenic plant seeds under conditions effective to impart disease resistance.

30

36. A method according to claim 35, wherein a transgenic plant is provided.

- 79 -

37. A method according to claim 35, wherein a transgenic plant seed is provided.

5 38. A method of enhancing plant growth comprising:
 providing a transgenic plant or a plant seed transformed with a DNA
 molecule which encodes a fragment of a hypersensitive response elicitor protein or
 polypeptide, which fragment elicits a hypersensitive response, and
 growing the transgenic plant or transgenic plants produced from the
10 transgenic plant seeds under conditions effective to enhance plant growth.

39. A method according to claim 38, wherein a transgenic plant is provided.

15 40. A method according to claim 38, wherein a transgenic plant seed is provided.

 41. A method of insect control for plants comprising:
 providing a transgenic plant or plant seed transformed with a DNA
20 molecule which encodes a fragment of a hypersensitive response elicitor protein or
 polypeptide, which fragment elicits a hypersensitive response, and
 growing the transgenic plant or transgenic plants produced from the
 transgenic plant seeds under conditions effective to control insects.

25 42. A method according to claim 41, wherein a transgenic plant is provided.

 43. A method according to claim 41, wherein a transgenic plant seed is provided.

30

1/13

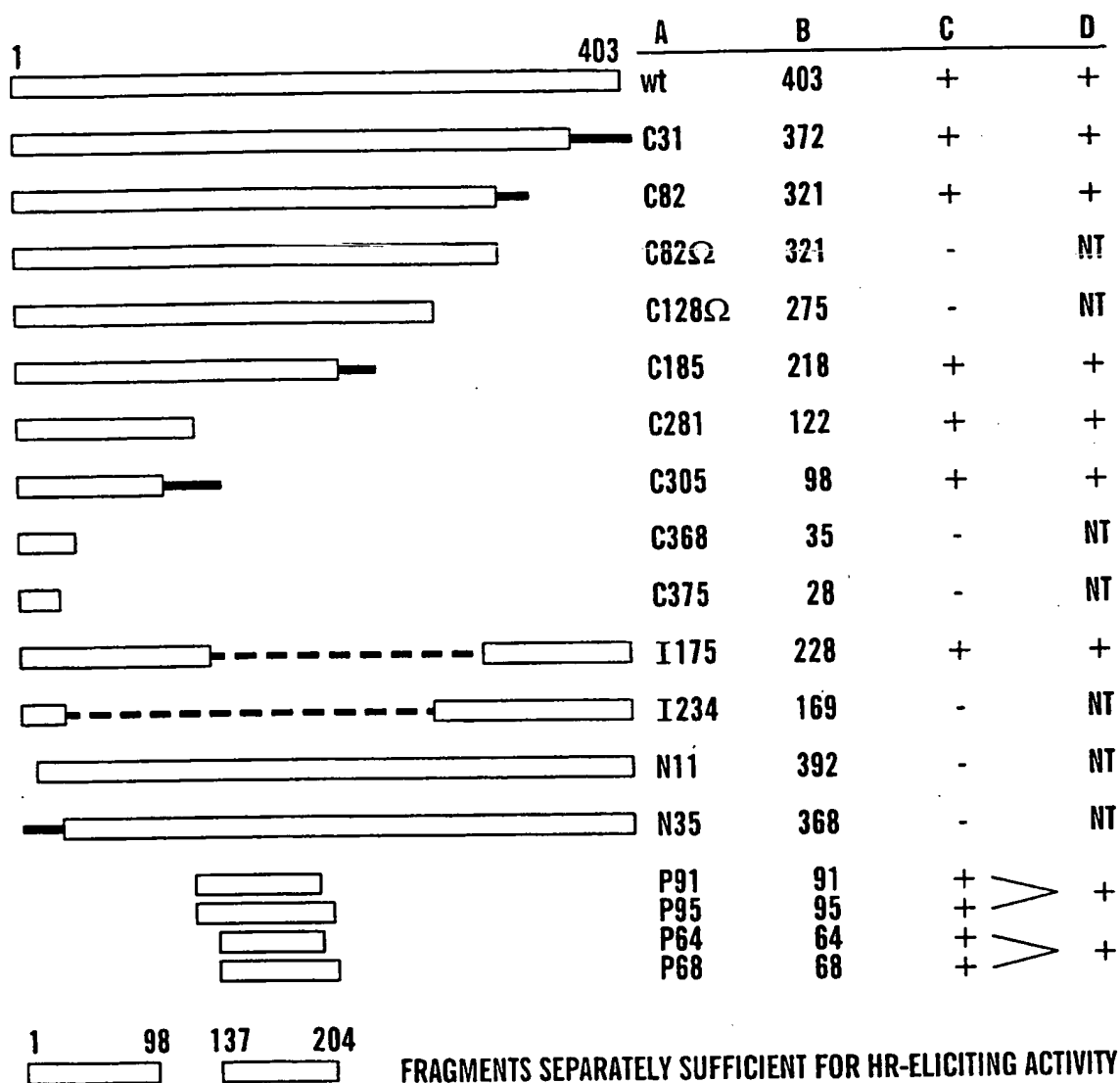


FIG. 1

2/13

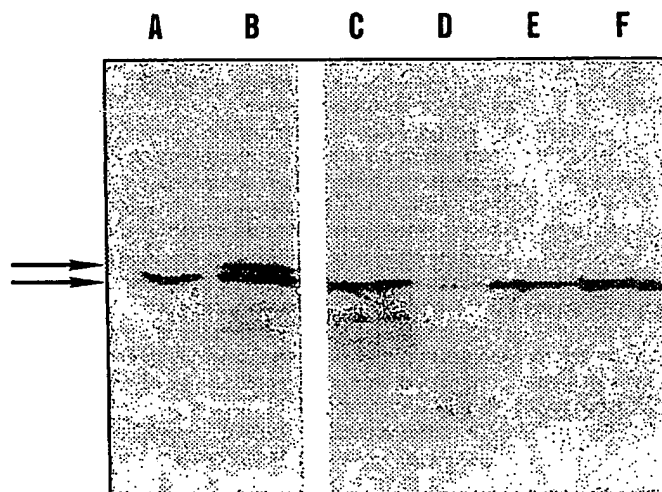


FIG. 2

3/13

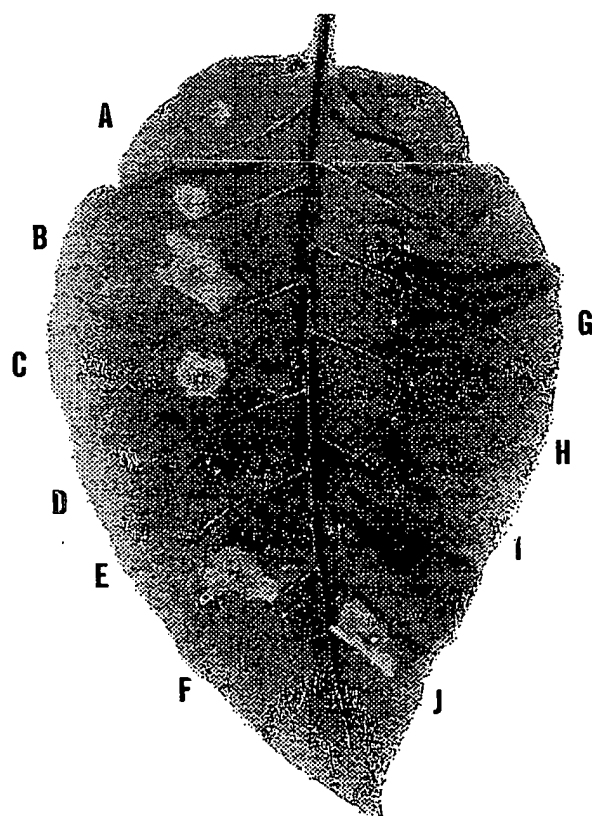


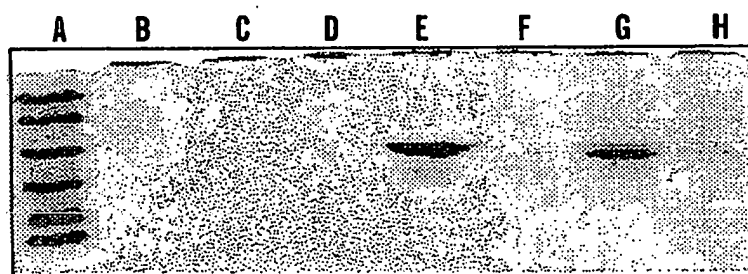
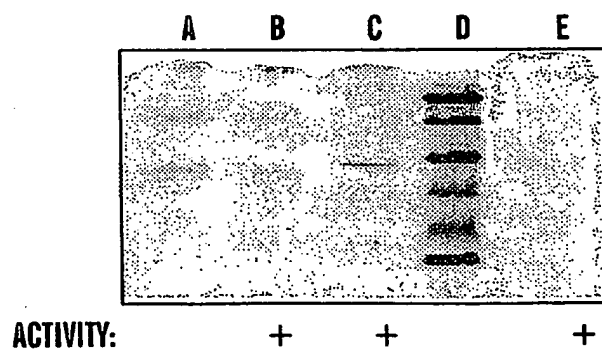
FIG. 3

4/13



FIG. 4

5/13

**FIG. 5A****FIG. 5B**

6/13

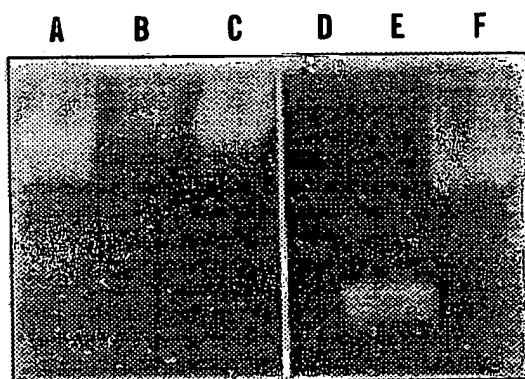
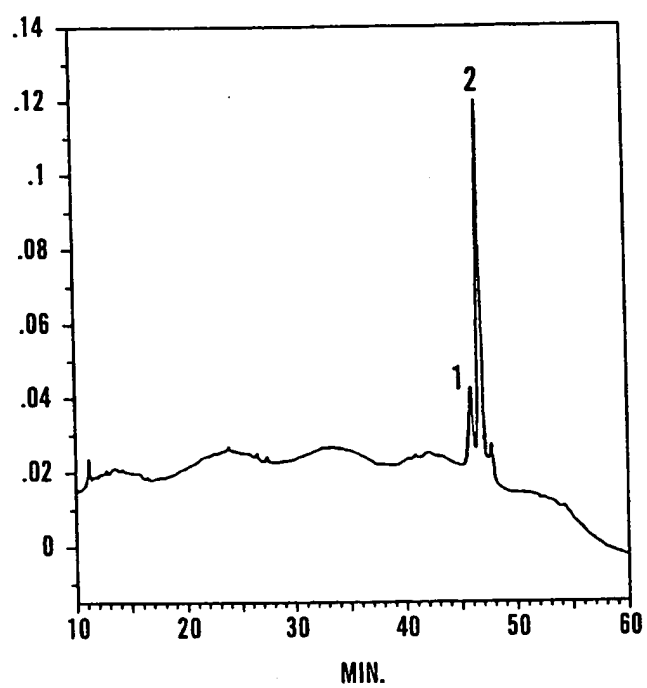


FIG. 5C

7/13

**FIG. 6**

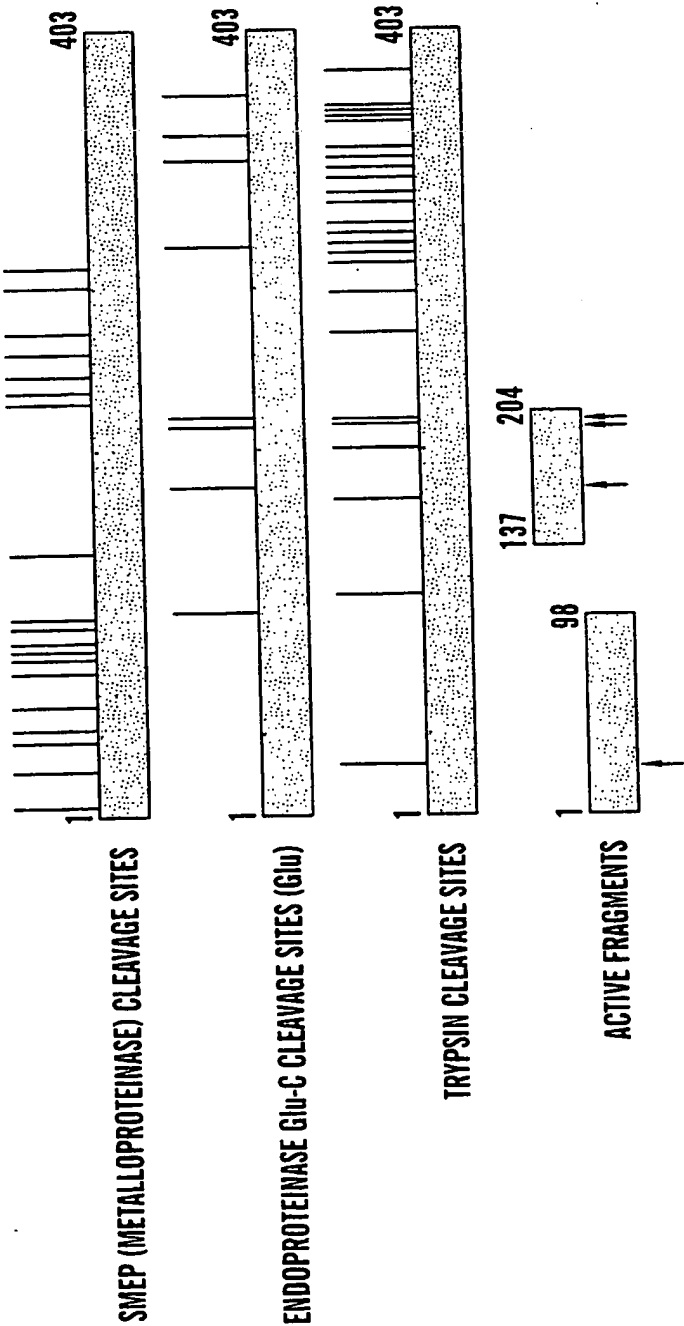


FIG. 7

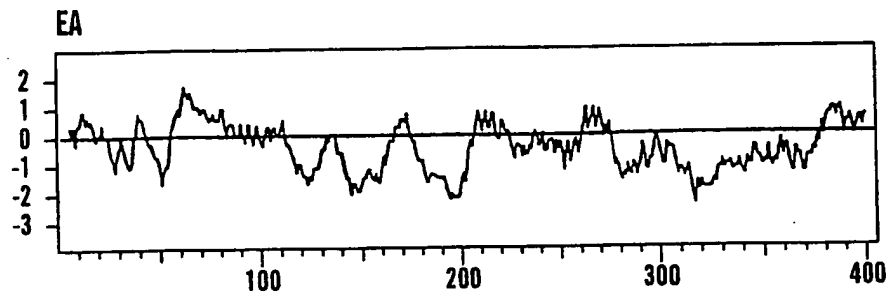
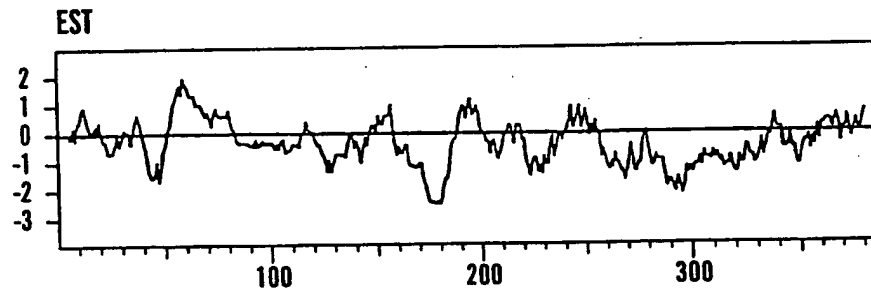
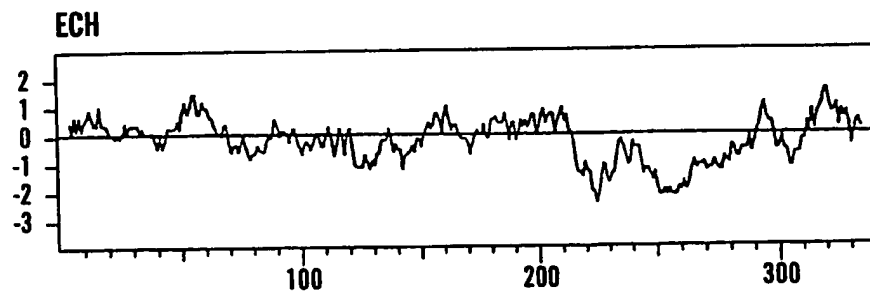
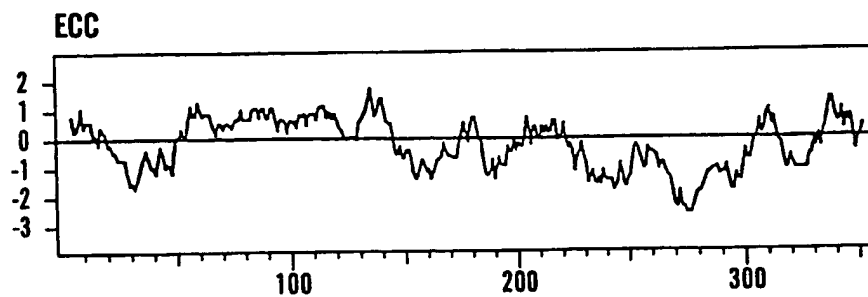
9/13

Gene	Organism ^a	N-terminal amino acid sequence
HrpN	Ech	MQITIKAHIGDDLGV <u>SLGAQGLKGL</u> NSAA
HrpN	Ecc	M.L <u>NSLGGAS</u> .LQITIK.AGGNGGLFPSQ
HrpN	Eam	MSLNT <u>SLGASTMQISTGGAGGNNGL</u> LGTS
WtsN	Est	MSMNT <u>SPPLGTSALQVTL</u> ...GGNNGLMGTD
orf1B	Eam	MSILTLNNNTSS.SPGLFQSGGD <u>NGLG</u> GHA

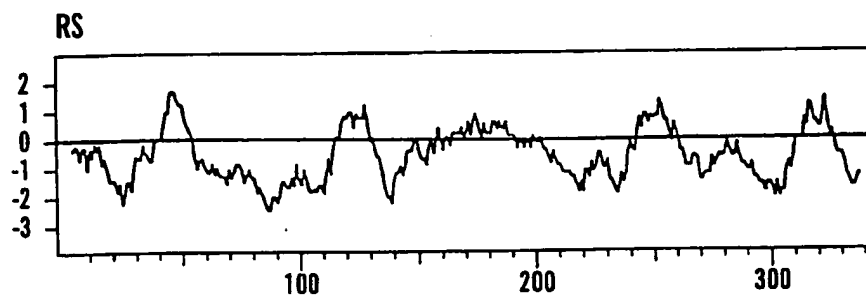
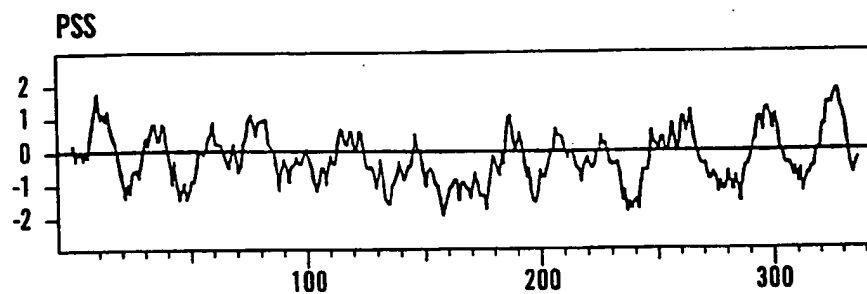
^aEch= *E. chrysanthemi*; Ecc= *E. carotovora* subsp. *carotovora*;
 Eam= *E. amylovora*; Est= *E. stewartii*.

FIG. 8

10/13

**FIG. 9A****FIG. 9B****FIG. 9C****FIG. 9D**

11/13

**FIG. 9E****FIG. 9F**

12/13

	HARPIN		HR
#1	1	403	+
	C-TERMINAL FRAGMENTS		
#3	105	403	+
#4	169	403	-
#5	210	403	-
#6	267	403	-
#7	343	403	-
	N-TERMINAL FRAGMENTS		
#8	1	75	-
#9	1	104	+/-
#10	1	168	+
#11	1	266	+
#12	1	342	+
	INTERNAL FRAGMENTS		
#13	76	209	+
#14	76	168	-
#15	105	209	+
#16	169	209	NA
#17	105	168	-
	SYNTHESIZED OLIGOPEPTIDES		
#18	99	209	+
#19	137	204	+
#20	137	180	+
#21	105	180	+
#22	150	209	NA
#23	150	180	NA

FIG. 10

13/13

N1; 5'-GGGAATTCATATGAGTCTGAATACAAGTGGG-3'
N76; 5'-GGGAATTCATATGGGCGGTGGCTTAGGCGGT-3'
N99; 5'-GGCATATGTCGAACGCGCTGAACGATATG-3'
N105; 5'-GGGAATTCATATGTTAGGCGGTTCGCTGAAC-3'
N110; 5'-GGCATATGCTGAACACGCTGGGCTCGAAA-3'
N137; 5'-GGCATATGTCAACGTCCCAAACGACGAT-3'
N150; 5'-GGCATATGTCCACCTCAGACTCCAGCG-3'
N169; 5'-GGGAATTCATATGCAAAGCCTGTTTGGTGATGGG-3'
N210; 5'-GGGAATTCATATGGGTAATGGTCTGAGCAAG-3'
N267; 5'-GGGAATTCATATGAAAGCGGGCATTTCAGGCG-3'
N343; 5'-GGGAATTCATATGACACCAGCCAGTATGGAGCAG-3'
C75; 5'-GCAAGCTTAACAGCCCACCACCGCCCATCAT-3'
C104; 5'-GCAAGCTTAAATCGTTCAGCGCGTTCGACAG-3'
C168; 5'-GCAAGCTTAAATATCTCGCTGAACATCTTCAGCAG-3'
C180; 5'-GCAAGCTTAAGGTGCCATCTTGCCCATCAC-3'
C204; 5'-GCAAGCTTAAATCAGTGACTCCTTTTTTATAGGC-3'
C209; 5'-GCAAGCTTAACAGGCCCGACAGCGCATCAGT-3'
C266; 5'-GCAAGCTTAAACCGATAACCGGTACCCACGGC-3'
C342; 5'-GCAAGCTTAATCCGTCGTCATCTGGCTTGCTCAG-3'
C403; 5'-GCAAGCTTAAGCCGCGCCCAGCTTG-3'

FIG. 11

INTERNATIONAL SEARCH REPORT

International Application No
PCT/US 98/10874

A. CLASSIFICATION OF SUBJECT MATTER IPC 6 C12N15/31 C07K14/27 A01H1/00 A01H3/00		
According to International Patent Classification (IPC) or to both national classification and IPC		
B. FIELDS SEARCHED Minimum documentation searched (classification system followed by classification symbols) IPC 6 C12N C07K		
Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched		
Electronic data base consulted during the international search (name of data base and, where practical, search terms used)		
C. DOCUMENTS CONSIDERED TO BE RELEVANT		
Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	WO 96 39802 A (CORNELL RES FOUNDATION INC) 19 December 1996 see abstract see page 10, line 30 - page 11, line 17 see page 12, line 21 - page 14, line 33 see page 17, line 33 - page 19, line 51 see page 27, line 11 - line 25 see page 29, line 28 - line 36 see examples 1-9 see figures 1,2 <div style="text-align: center;">--- -/--</div>	1-3, 8-10, 15-28, 35-37
<div style="display: flex; justify-content: space-between;"> <input checked="" type="checkbox"/> Further documents are listed in the continuation of box C. <input checked="" type="checkbox"/> Patent family members are listed in annex. </div>		
* Special categories of cited documents : <div style="display: flex;"> <div style="flex: 1;"> <p>"A" document defining the general state of the art which is not considered to be of particular relevance</p> <p>"E" earlier document but published on or after the international filing date</p> <p>"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)</p> <p>"O" document referring to an oral disclosure, use, exhibition or other means</p> <p>"P" document published prior to the international filing date but later than the priority date claimed</p> </div> <div style="flex: 1;"> <p>"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention</p> <p>"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone</p> <p>"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.</p> <p>"&" document member of the same patent family</p> </div> </div>		
Date of the actual completion of the international search <div style="text-align: center;">20 November 1998</div>		Date of mailing of the international search report <div style="text-align: center;">28/12/1998</div>
Name and mailing address of the ISA European Patent Office, P.B. 5818 Patentlaan 2 NL - 2280 HV Rijswijk Tel. (+31-70) 340-2040, Tx. 31 651 epo nl, Fax: (+31-70) 340-3016		Authorized officer <div style="text-align: center;">Panzica, G</div>

INTERNATIONAL SEARCH REPORT

International Application No
PCT/US 98/10874

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	EP 0 612 848 A (SANDOZ AG ; SANDOZ LTD (CH); SANDOZ AG (DE)) 31 August 1994 see page 2, line 28 - line 35 see page 2, line 51 - line 54 see page 4, line 23 - line 30	1-3, 8-10
A	see page 5, line 12 - line 18	4-7, 11-43
X	WO 94 26782 A (CORNELL RES FOUNDATION INC) 24 November 1994 see page 2, line 11 - line 25 see page 3, line 14 - line 28	1-3, 8-10, 26-28
A		4-7, 11-14
X	WO 94 01546 A (CORNELL RES FOUNDATION INC) 20 January 1994 see abstract see page 3, line 9 - page 4, line 10 see examples 1-15	1-3, 8-10, 15-28, 35-37
A		4-7, 11-14
A	WO 93 23532 A (RIBOZYME PHARM INC) 25 November 1993 see abstract see page 1, line 23 - page 2, line 8 see page 2, line 25 - page 3, line 5	26-28, 35-37
X	WEI Z.M. ET AL.: "Harpin, elicitor of the hypersensitive response produced by the plant pathogen Erwinia amylovora" SCIENCE, vol. 257, 3 July 1992, pages 85-88, XP002085274 LANCASTER, PA US cited in the application see the whole document	1-3, 8
X	BAUER D.W. ET AL.: "Erwinia chrysanthemi Harpin(Ech): an elicitor of the hypersensitive response that contributes to soft-rot pathogenesis" MOLECULAR PLANT-MICROBE INTERACTIONS, vol. 8, no. 4, 1995, pages 484-491, XP002085275 cited in the application see the whole document	1-3, 8
A		4-7, 11-14
	-/--	

INTERNATIONAL SEARCH REPORT

International Application No
PCT/US 98/10874

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT		
Category	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	WEI Z.M., BEER S.V.: "hrpL activates Erwinia amylovora hrp gene transcription and is a member of the ECF subfamily of "sigma" factors" JOURNAL OF BACTERIOLOGY, vol. 177, no. 21, November 1995, pages 6201-6210, XP002085276 see abstract	1-3,8
A	see figure 7 -----	4-7, 11-14

INTERNATIONAL SEARCH REPORT

Information on patent family members

In International Application No

PCT/US 98/10874

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
WO 9639802 A	19-12-1996	US 5650387 A	22-07-1997
		AU 5982196 A	30-12-1996
		CA 2223616 A	19-12-1996
		CN 1192647 A	09-09-1998
		EP 0871354 A	21-10-1998
		FI 974430 A	26-01-1998
		PL 323823 A	27-04-1998
		US 5776889 A	07-07-1998
EP 0612848 A	31-08-1994	AU 682674 B	16-10-1997
		AU 5631194 A	01-09-1994
		CA 2116202 A	25-08-1994
		CZ 9400404 A	12-07-1995
		HU 69614 A	28-09-1995
		JP 6319558 A	22-11-1994
		PL 302320 A	05-09-1994
		SK 20694 A	10-05-1995
		US 5550228 A	27-08-1996
		US 5552527 A	03-09-1996
		ZA 9401281 A	24-08-1995
WO 9426782 A	24-11-1994	US 5708139 A	13-01-1998
		EP 0702694 A	27-03-1996
		JP 8510127 T	29-10-1996
WO 9401546 A	20-01-1994	EP 0648266 A	19-04-1995
		JP 7509604 T	26-10-1995
WO 9323532 A	25-11-1993	AU 4234293 A	13-12-1993
		CA 2135643 A	25-11-1993
		EP 0641384 A	08-03-1995
		JP 7507931 T	07-09-1995
		MX 9302804 A	29-04-1994



INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification ⁶ : A01G 13/00, A61K 35/66, C12N 1/20, C12R 1/18	A1	(11) International Publication Number: WO 99/07206 (43) International Publication Date: 18 February 1999 (18.02.99)
(21) International Application Number: PCT/US98/15426 (22) International Filing Date: 24 July 1998 (24.07.98) (30) Priority Data: 60/055,105 6 August 1997 (06.08.97) US (71) Applicant: CORNELL RESEARCH FOUNDATION, INC. [US/US]; Suite 105, 20 Thornwood Drive, Ithaca, NY 14850 (US). (72) Inventors: BOGDANOVE, Adam, J.; 210-2 Airport Road, West Lafayette, IN 47906 (US). KIM, Jihyun, Francis; Apartment E-4, 2250 N. Triphammer Road (US). WEI, Zhong-Min; 8230 125th Court, Kirkland, WA 98034 (US). BEER, Steven, V.; 211 Hudson Street, Ithaca, NY 14850 (US). (74) Agents: GOLDMAN, Michael, L. et al.; Nixon, Hargrave, Devans & Doyle LLP, Clinton Square, P.O. Box 1051, Rochester, NY 14603 (US).		(81) Designated States: AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE, DK, EE, ES, FI, GB, GE, GH, GM, HU, ID, IL, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, UA, UG, UZ, VN, YU, ZW, ARIPO patent (GH, GM, KE, LS, MW, SD, SZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG). Published <i>With international search report.</i>
(54) Title: HYPERSENSITIVE RESPONSE ELICITOR FROM <i>ERWINIA AMYLOVORA</i> , ITS USE, AND ENCODING GENE (57) Abstract <p>The present invention is directed to an isolated protein or polypeptide which elicits a hypersensitive response in plants as well as an isolated DNA molecule which encodes the hypersensitive response eliciting protein or polypeptide. This isolated protein or polypeptide and the isolated DNA molecule can be used to impart disease resistance to plants, to enhance plant growth, and/or to control insects on plants. This can be achieved by applying the hypersensitive response elicitor protein or polypeptide in a non-infectious form to plants or plant seeds under conditions effective to impart disease resistance, to enhance plant growth, and/or to control insects on plants or plants grown from the plant seeds. Alternatively, transgenic plants or plant seeds transformed with a DNA molecule encoding a hypersensitive response elicitor protein or polypeptide can be provided and the transgenic plants or plants resulting from the transgenic plant seeds are grown under conditions effective to impart disease resistance, to enhance plant growth, and/or to control insects on plants or plants grown from the plant seeds.</p>		

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	ES	Spain	LS	Lesotho	SI	Slovenia
AM	Armenia	FI	Finland	LT	Lithuania	SK	Slovakia
AT	Austria	FR	France	LU	Luxembourg	SN	Senegal
AU	Australia	GA	Gabon	LV	Latvia	SZ	Swaziland
AZ	Azerbaijan	GB	United Kingdom	MC	Monaco	TD	Chad
BA	Bosnia and Herzegovina	GE	Georgia	MD	Republic of Moldova	TG	Togo
BB	Barbados	GH	Ghana	MG	Madagascar	TJ	Tajikistan
BE	Belgium	GN	Guinea	MK	The former Yugoslav Republic of Macedonia	TM	Turkmenistan
BF	Burkina Faso	GR	Greece	ML	Mali	TR	Turkey
BG	Bulgaria	HU	Hungary	MN	Mongolia	TT	Trinidad and Tobago
BJ	Benin	IE	Ireland	MR	Mauritania	UA	Ukraine
BR	Brazil	IL	Israel	MW	Malawi	UG	Uganda
BY	Belarus	IS	Iceland	MX	Mexico	US	United States of America
CA	Canada	IT	Italy	NE	Niger	UZ	Uzbekistan
CF	Central African Republic	JP	Japan	NL	Netherlands	VN	Viet Nam
CG	Congo	KE	Kenya	NO	Norway	YU	Yugoslavia
CH	Switzerland	KG	Kyrgyzstan	NZ	New Zealand	ZW	Zimbabwe
CI	Côte d'Ivoire	KP	Democratic People's Republic of Korea	PL	Poland		
CM	Cameroon	KR	Republic of Korea	PT	Portugal		
CN	China	KZ	Kazakhstan	RO	Romania		
CU	Cuba	LC	Saint Lucia	RU	Russian Federation		
CZ	Czech Republic	LI	Liechtenstein	SD	Sudan		
DE	Germany	LK	Sri Lanka	SE	Sweden		
DK	Denmark	LR	Liberia	SG	Singapore		
EE	Estonia						

- 1 -

**HYPERSENSITIVE RESPONSE ELICITOR FROM
ERWINIA AMYLOVORA, ITS USE, AND ENCODING GENE**

This application claims benefit of U.S. Provisional Patent Application
5 Serial No. 60/055,105, filed August 4, 1997.

FIELD OF THE INVENTION

The present invention relates to a hypersensitive response elicitor from
10 *Erwinia amylovora*, its use, and encoding gene.

BACKGROUND OF THE INVENTION

Interactions between bacterial pathogens and their plant hosts generally
15 fall into two categories: (1) compatible (pathogen-host), leading to intercellular
bacterial growth, symptom development, and disease development in the host plant;
and (2) incompatible (pathogen-nonhost), resulting in the hypersensitive response, a
particular type of incompatible interaction occurring, without progressive disease
symptoms. During compatible interactions on host plants, bacterial populations
20 increase dramatically and progressive symptoms occur. During incompatible
interactions, bacterial populations do not increase, and progressive symptoms do not
occur.

The hypersensitive response ("HR") is a rapid, localized necrosis that
is associated with the active defense of plants against many pathogens (Kiraly, Z.,
25 "Defenses Triggered by the Invader: Hypersensitivity," pages 201-224 in: Plant
Disease: An Advanced Treatise, Vol. 5, J.G. Horsfall and E.B. Cowling, ed.
Academic Press New York (1980); Klement, Z., "Hypersensitivity," pages 149-177
in: Phytopathogenic Prokaryotes, Vol. 2, M.S. Mount and G.H. Lacy, ed. Academic
Press, New York (1982)). The hypersensitive response elicited by bacteria is readily
30 observed as a tissue collapse if high concentrations ($\geq 10^7$ cells/ml) of a limited
host-range pathogen like *Pseudomonas syringae* or *Erwinia amylovora* are infiltrated
into the leaves of nonhost plants (necrosis occurs only in isolated plant cells at lower
levels of inoculum) (Klement, Z., "Rapid Detection of Pathogenicity of
Phytopathogenic Pseudomonads," Nature 199:299-300; Klement, et al.,

- 2 -

“Hypersensitive Reaction Induced by Phytopathogenic Bacteria in the Tobacco Leaf,” Phytopathology 54:474-477 (1963); Turner, et al., “The Quantitative Relation Between Plant and Bacterial Cells Involved in the Hypersensitive Reaction,” Phytopathology 64:885-890 (1974); Klement, Z., “Hypersensitivity,” pages 149-177 in Phytopathogenic Prokaryotes, Vol. 2., M.S. Mount and G.H. Lacy, ed. Academic Press, New York (1982)). The capacities to elicit the hypersensitive response in a nonhost and be pathogenic in a host appear linked. As noted by Klement, Z., “Hypersensitivity,” pages 149-177 in Phytopathogenic Prokaryotes, Vol. 2., M.S. Mount and G.H. Lacy, ed. Academic Press, New York, these pathogens also cause physiologically similar, albeit delayed, necroses in their interactions with compatible hosts. Furthermore, the ability to produce the hypersensitive response or pathogenesis is dependent on a common set of genes, denoted *hrp* (Lindgren, P.B., et al., “Gene Cluster of *Pseudomonas syringae* pv. ‘phaseolicola’ Controls Pathogenicity of Bean Plants and Hypersensitivity on Nonhost Plants,” J. Bacteriol. 168:512-22 (1986); Willis, D.K., et al., “*hrp* Genes of Phytopathogenic Bacteria,” Mol. Plant-Microbe Interact. 4:132-138 (1991)). Consequently, the hypersensitive response may hold clues to both the nature of plant defense and the basis for bacterial pathogenicity.

The *hrp* genes are widespread in gram-negative plant pathogens, where they are clustered, conserved, and in some cases interchangeable (Willis, D.K., et al., “*hrp* Genes of Phytopathogenic Bacteria,” Mol. Plant-Microbe Interact. 4:132-138 (1991); Bonas, U., “*hrp* Genes of Phytopathogenic Bacteria,” pages 79-98 in: Current Topics in Microbiology and Immunology: Bacterial Pathogenesis of Plants and Animals - Molecular and Cellular Mechanisms, J.L. Dangel, ed. Springer-Verlag, Berlin (1994)). Several *hrp* genes encode components of a protein secretion pathway similar to one used by *Yersinia*, *Shigella*, and *Salmonella* spp. to secrete proteins essential in animal diseases (Van Gijsegem, et al., “Evolutionary Conservation of Pathogenicity Determinants Among Plant and Animal Pathogenic Bacteria,” Trends Microbiol. 1:175-180 (1993)). In *E. amylovora*, *P. syringae*, and *P. solanacearum*, *hrp* genes have been shown to control the production and secretion of glycine-rich, protein elicitors of the hypersensitive response (He, S.Y., et al. “*Pseudomonas Syringae* pv. *Syringae* Harpin_{PS}: a Protein that is Secreted via the Hrp Pathway and Elicits the Hypersensitive Response in Plants,” Cell 73:1255-1266 (1993). Wei, Z.-H.,

- 3 -

et al., "HrpI of *Erwinia amylovora* Functions in Secretion of Harpin and is a Member of a New Protein Family," J. Bacteriol. 175:7958-7967 (1993); Arlat, M. et al. "PopA1, a Protein Which Induces a Hypersensitive-like Response on Specific Petunia Genotypes, is Secreted via the Hrp Pathway of *Pseudomonas solanacearum*," EMBO J. 13:543-553 (1994)).

The first of these proteins was discovered in *E. amylovora* Ea321, a bacterium that causes fire blight of rosaceous plants, and was designated harpin (Wei, Z.-M., et al, "Harpin, Elicitor of the Hypersensitive Response Produced by the Plant Pathogen *Erwinia amylovora*," Science 257:85-88 (1992)). Mutations in the encoding *hrpN* gene revealed that the hypersensitive response elicitor is required for *E. amylovora* to elicit a hypersensitive response in nonhost tobacco leaves and incite disease symptoms in highly susceptible pear fruit. The *P. solanacearum* GM11000 PopA1 protein has similar physical properties and also elicits the hypersensitive response in leaves of tobacco, which is not a host of that strain (Arlat, et al. "PopA1, a Protein Which Induces a Hypersensitive-like Response on Specific Petunia Genotypes, is Secreted via the Hrp Pathway of *Pseudomonas solanacearum*," EMBO J. 13:543-53 (1994)). However, *P. solanacearum popA* mutants still elicit the hypersensitive response in tobacco and incite disease in tomato. Thus, the role of these glycine-rich hypersensitive response elicitors can vary widely among gram-negative plant pathogens.

Other plant pathogenic hypersensitive response elicitors have been isolated and their encoding genes have been cloned and sequenced. These include: *Erwinia chrysanthemi* (Bauer, et. al., "*Erwinia chrysanthemi* Harpin_{Ech}: Soft-Rot Pathogenesis," MPMI 8(4): 484-91 (1995)); *Erwinia carotovora* (Cui, et. al., "The RsmA⁻ Mutants of *Erwinia carotovora* subsp. *carotovora* Strain Ecc71 Overexpress *hrpN*_{Ecc} and Elicit a Hypersensitive Reaction-like Response in Tobacco Leaves," MPMI 9(7): 565-73 (1996)); *Erwinia stewartii* (Ahmad, et. al., "Harpin is not Necessary for the Pathogenicity of *Erwinia stewartii* on Maize," 8th Int'l. Cong. Molec. Plant-Microb. Inter. July 14-19, 1996 and Ahmad, et. al., "Harpin is not Necessary for the Pathogenicity of *Erwinia stewartii* on Maize," Ann. Mtg. Am. Phytopath. Soc. July 27-31, 1996); and *Pseudomonas syringae* pv. *syringae* (WO 94/26782 to Cornell Research Foundation, Inc.).

- 4 -

The present invention is a further advance in the effort to identify, clone, and sequence hypersensitive response elicitor proteins or polypeptides from plant pathogens.

5

SUMMARY OF THE INVENTION

The present invention is directed to an isolated protein or polypeptide which elicits a hypersensitive response in plants as well as an isolated DNA molecule which encodes the hypersensitive response eliciting protein or polypeptide.

10

The hypersensitive response eliciting protein or polypeptide can be used to impart disease resistance to plants, to enhance plant growth, and/or to control insects. This involves applying the hypersensitive response elicitor protein or polypeptide in a non-infectious form to plants or plant seeds under conditions effective to impart disease resistance, to enhance plant growth, and/or to control insects on plants or plants grown from the plant seeds.

15

As an alternative to applying the hypersensitive response elicitor protein or polypeptide to plants or plant seeds in order to impart disease resistance, to enhance plant growth, and/or to control insects on plants, transgenic plants or plant seeds can be utilized. When utilizing transgenic plants, this involves providing a transgenic plant transformed with a DNA molecule encoding a hypersensitive response elicitor protein or polypeptide and growing the plant under conditions effective to impart disease resistance, to enhance plant growth, and/or to control insects in the plants or plants grown from the plant seeds. Alternatively, a transgenic plant seed transformed with the DNA molecule encoding a hypersensitive response elicitor protein or polypeptide can be provided and planted in soil. A plant is then propagated under conditions effective to impart disease resistance, to enhance plant growth, and/or to control insects on plants or plants grown from the plant seeds.

20

25

BRIEF DESCRIPTION OF THE DRAWINGS

30

Figures 1A-D show mutagenesis, complementation and heterologous expression constructs, and homology with and complementation of mutants by the *avrE* locus of *P. syringae* for the *dspE* operon of *E. amylovora*. Dashed boxes are

- 5 -

uncharacterized ORFs; a filled triangle indicates a *hrp* (i.e. hypersensitive response elicitor encoding gene); box is a regulatory sequence that preceeds many *hrp* genes; and an open triangle indicates another promoter. Thick lines delineate the DNA for which sequence was accessioned. Figure 1A shows the *dsp/hrp* gene cluster of

5 *E. amylovora* in pCPP430. Operon names and types of proteins encoded are indicated at the top. B, *Bam*HI; E, *Eco*RI; H, *Hind*III. Half-arrows indicate internal promoters without similarity to the *hrp* box consensus. Figure 1B shows the region downstream of *hrpN* containing the *dspE* operon. Circles mark deletion mutations and representative transposon insertions: black, non-pathogenic and HR⁺ (i.e.

10 hypersensitive response eliciting) or HR reduced (*dsp*); gray, reduced virulence and HR; white, wild-type. T104 lies within the area marked by the dashed double arrow. K, *Tn5miniKm*; P, *Tn5phoA*; T, *Tn10tet*^r; Δ, deletion mutation. The gray box is A/T-rich DNA. Figure 1C shows the clones and subclones of the *dspE* operon. Plasmid designations are indicated at the left, and vector-borne promoters at the right.

15 Restriction sites used for subcloning not shown above are shown in parentheses. A "+" aligned with a circle representing a mutation in B indicates that the subclone complements that mutation. Figure 1D shows the *avrE* locus (transcription units III and IV) of *P. syringae* pv. tomato DC3000 in pCPP2357. Percent amino acid identity of the predicted proteins AvrE and AvrF to DspE and DspF, respectively, are

20 indicated. Black rectangles are transcriptional terminators (inverted repeats). Complementation of mutations shown in Figure 1B are depicted as in Figure 1C, above.

Figure 2 shows the expression of the full-length and the N-terminal half of DspE in recombinant *E. coli* DH5α. Lysates of cells carrying either

25 pCPP1259, containing the entire *dspE* operon (lane A); pCPP50, the cloning vector (lane B); or pCPP1244, containing only the 5' half of the *dspE* gene (lane C), were subjected to SDS-PAGE followed by Coomassie staining. Bands corresponding to DspE (lane A) and the N-terminal half of DspE (lane C) are marked by arrows. Migration of molecular weight markers is indicated on the left.

30 Figures 3A-D show the role of *dspE* in pathogenicity and HR elicitation. Figure 3A shows immature pear fruit 4 days after inoculation with (left to right) strains Ea321, Ea321*dspE*Δ1554, or Ea321*dspE*Δ1554 harboring the 5' half of

- 6 -

dspE on pCPP1237. Figure 3B shows Norchief soybean leaf 24 hr after infiltration with (1) Ea321, (2) Ea321*dspE*Δ1554, (3) Ea321*hrpN::Tn5* (Wei, et al., Science, 257:85-88 (1992), which is hereby incorporated by reference), and (4) Ea321*hrpL::Tn5* (Wei, et al., J. Bacteriol., 177:6201-10 (1995), which is hereby incorporated by reference). Figure 3C shows a tobacco leaf 48 hr after infiltration with parallel dilution series of suspensions of strains (left) Ea321 and (right) Ea321*dspE*Δ1554. The concentrations infiltrated (top to bottom) are 1×10^{10} , 1×10^9 , 5×10^8 , and 5×10^7 cfu/ml. Figure 3D is like Figure 3C except the more virulent strain Ea273 and corresponding mutant Ea273*dspE*Δ1554 were used, and concentrations ranged from 5×10^9 to 5×10^5 cfu/ml in log increments.

Figure 4 shows the expression of a promoterless GUS construct fused to *dspE* in *E. amylovora* Ea273. Ea273 and Ea273*dspE::uidA* (a merodiploid containing both a wild-type *dspE* and a truncated *dspE* fused to the *uidA* gene; black bars) were grown in LB or Hrp MM, or inoculated to immature pear fruit. Ea273*dspE::uidAhrpL::Tn5* (dark gray bar) and Ea273*hrcV::Tn5uidA* (light gray bar) were also grown in *hrp* MM. Values shown represent means of triplicate samples normalized for bacterial cell concentration. Standard deviations are represented by lines extending from each bar. The mean values for three samples of Ea273 in each assay were subtracted from, and standard deviations added to, the corresponding values obtained for the other strains.

Figures 5A-C show the transgeneric avirulence function of the *dspE* operon and complementation of a *dspE* mutant with the *avrE* locus. Norchief soybean leaves were either (See Figure 5A) infiltrated with 1×10^8 cfu/ml suspensions of (left) *P. syringae* pv. *glycinea* race 4 carrying pCPP1250 (containing the *dspE* operon) or (right) pML 122 (the cloning vector) and photographed after 24 hr at room temperature or (See Figure 5B) infiltrated with 8×10^5 cfu/ml suspensions of the same strains and photographed after seven days at 22° C and high relative humidity. Tissue collapse is apparent on both leaves where the strain carrying pCPP1250 was infiltrated. On the leaf incubated for seven days, chlorosis extending beyond the infiltrated area, typical of disease, is apparent on the half infiltrated with the strain carrying the vector only. The dark section on the side of the leaf infiltrated with the strain carrying pCPP1250 is a shadow caused by a buckle in the leaf. Figure 5C

- 7 -

shows pear halves inoculated with (left to right) Ea321, Ea321*dspE*Δ1521(pCPP2357, containing the *avrE* locus), or Ea321*dspE*Δ1521(pCPP2357*avrE::Tn5uidA*) and photographed after seven days. Although symptoms are greatly reduced relative to wild type, necrosis is apparent around and drops of ooze can be seen within the well of the fruit inoculated with the *dspE* strain carrying the intact *avrE* locus. The fruit inoculated with the *dspE* strain carrying a disrupted clone of *avrE* is symptomless.

DETAILED DESCRIPTION OF THE INVENTION

The present invention relates to an isolated DNA molecule having a nucleotide sequence of SEQ. ID. No. 1 as follows:

	ATGGAATTAA AATCACTGGG AACTGAACAC AAGGCGGCAG TACACACAGC GGCGCACAAAC	60
15	CCTGTGGGGC ATGGTGTTC CTTACAGCAG GGCAGCAGCA GCAGCAGCCC GCAAAATGCC	120
	GCTGCATCAT TGGCGGCAGA AGGCAAAAAT CGTGGGAAAA TGCCGAGAAT TCACCAGCCA	180
20	TCTACTGCGG CTGATGGTAT CAGCGCTGCT CACCAGCAAA AGAAATCCTT CAGTCTCAGG	240
	GGCTGTITGG GGACGAAAAA ATTTTCCAGA TCGGCACCGC AGGGCCAGCC AGGTACCACC	300
	CACAGCAAAG GGGCAACATT GCGCGATCTG CTGGCGCGGG ACGACGGCGA AACGCAGCAT	360
25	GAGGCGGCCG CGCCAGATGC GGCGCGTTTG ACCCGTTCGG GCGGCGTCAA ACGCCGCAAT	420
	ATGGACGACA TGGCCGGGCG GCCAATGGTG AAAGGTGGCA GCGGCGAAGA TAAGGTACCA	480
30	ACGCAGCAAA AACGGCATCA GCTGAACAAT TTTGGCCAGA TGCGCCAAAC GATGTTGAGC	540
	AAAATGGCTC ACCCGGCTTC AGCCAACGCC GGCGATCGCC TGCAGCATTC ACCGCCGCAC	600
	ATCCCGGGTA GCCACCACGA AATCAAGGAA GAACCGGTTG GCTCCACCAG CAAGGCAACA	660
35	ACGGCCACG CAGACAGAGT GGAAATCGCT CAGGAAGATG ACGACAGCGA ATTCCAGCAA	720
	CTGCATCAAC AGCGGCTGGC GCGCGAACGG GAAAATCCAC CGCAGCCGCC CAAACTCGGC	780
40	GTTGCCACAC CGATTAGCGC CAGGTTTCAG CCCAACTGA CTGCGGTTGC GGAAAGCGTC	840
	CTTGAGGGGA CAGATACCAC GCAGTCACCC CTTAAGCCGC AATCAATGCT GAAAGGAAGT	900
	GGAGCCGGG TAACGCCGCT GCGGTAACG CTGGATAAAG GCAAGTTGCA GCTGGCACCG	960
45	GATAATCCAC CCGCGCTCAA TACGTTGTTG AAGCAGACAT TGGGTAAAGA CACCCAGCAC	1020
	TATCTGGCGC ACCATGCCAG CAGCGACGGT AGCCAGCATC TGCTGCTGGA CAACAAAGGC	1080
50	CACCTGTTTG ATATCAAAAG CACCGCCACC AGCTATAGCG TGCTGCACAA CAGCCACCCC	1140
	GGTGAGATAA AGGGCAAGCT GGCGCAGGCG GGTACTGGCT CCGTCAGCGT AGACGGTAAA	1200

- 8 -

	AGCGGCAAGA TCTCGCTGGG GAGCGGTACG CAAAGTCACA AAAAAACAAT GCTAAGCCAA	1260
	CCGGGGGAAG CGCACCGTTC CTTATTAACC GGCATTGGC AGCATCCTGC TGGCGCAGCG	1320
5	CGGCCGCAGG GCGAGTCAAT CCGCCTGCAT GACGACAAAA TTCATATCCT GCATCCGGAG	1380
	CTGGGCGTAT GGCAATCTGC GGATAAAGAT ACCCACAGCC AGCTGTCTCG CCAGGCAGAC	1440
10	GGTAAGCTCT ATGCGCTGAA AGACAACCGT ACCCTGCAAA ACCTCTCCGA TAATAAATCC	1500
	TCAGAAAAGC TGGTCGATAA AATCAAATCG TATTCCGTTG ATCAGCGGGG GCAGGTGGCG	1560
	ATCCTGACGG ATACTCCCGG CCGCCATAAG ATGAGTATTA TGCCCTCGCT GGATGCTTCC	1620
15	CCGGAGAGCC ATATTTCCCT CAGCCTGCAT TTTGCCGATG CCCACCAGGG GTTATTGCAC	1680
	GGGAAGTCGG AGCTTGAGGC ACAATCTGTC GCGATCAGCC ATGGGCGACT GGTGTGGCC	1740
20	GATAGCGAAG GCAAGCTGTT TAGCGCCGCC ATTCCGAAGC AAGGGGATGG AAACGAACTG	1800
	AAAATGAAAG CCATGCCTCA GCATGCGCTC GATGAACATT TTGGTCATGA CCACCAGATT	1860
	TCTGGATTTT TCCATGACGA CCACGCCAG CTTAATGCGC TGGTGAAAAA TAACTTCAGG	1920
25	CAGCAGCATG CCTGCCCGTT GGGTAACGAT CATCAGTTTC ACCCCGGCTG GAACCTGACT	1980
	GATGCGCTGG TTATCGACAA TCAGCTGGGG CTGCATCATA CCAATCCTGA ACCGCATGAG	2040
30	ATTCITGATA TGGGGCATT T AGGCAGCCTG GCGTTACAGG AGGGCAAGCT TCACTATTTT	2100
	GACCAGCTGA CCAAAGGGTG GACTGGCGCG GAGTCAGATT GTAAGCAGCT GAAAAAAGGC	2160
	CTGGATGGAG CAGCTTATCT ACTGAAAGAC GGTGAAGTGA AACGCCTGAA TATTAATCAG	2220
35	AGCACCTCCT CTATCAAGCA CGGAACGGAA AACGTTTTTT CGCTGCCGCA TGTGCGCAAT	2280
	AAACCGGAGC CGGGAGATGC CCTGCAAGGG CTGAATAAAG ACGATAAGGC CCAGGCCATG	2340
40	GCGGTGATTG GGGTAAATAA ATACCTGGCG CTGACGAAA AAGGGGACAT TCGCTCCTTC	2400
	CAGATAAAAC CCGGCACCCA GCAGTTGGAG CGGCCGGCAC AACTCTCAG CCGCGAAGGT	2460
	ATCAGCGCG AACTGAAAGA CATTATGTC GACCACAAGC AGAACCTGTA TGCCTTGACC	2520
45	CACGAGGGAG AGGTGTTTCA TCAGCCGCGT GAAGCCTGGC AGAATGGTGC CGAAAGCAGC	2580
	AGCTGGCACA AACTGGCGTT GCCACAGAGT GAAAGTAAGC TAAAAAGTCT GGACATGAGC	2640
50	CATGAGCACA AACCGATTGC CACCTTTGAA GACGGTAGCC AGCATCAGCT GAAGGCTGGC	2700
	GGCTGGCAGC CCTATGCGGC ACCTGAAAGC GGGCCGCTGG CGGTGGGTAC CAGCGGTTCA	2760
	CAAACCGTCT TTAACCGACT AATGCAGGGG GTGAAAGGCA AGGTGATCCC AGGCAGCGGG	2820
55	TTGACGGTTA AGCTCTCGGC TCAGACGGGG GGAATGACCG GCGCCGAAGG GCGCAAGGTC	2880
	AGCAGTAAAT TTTCCGAAAG GATCCGCGCC TATGCGTTCA ACCCAACAAT GTCCACGCCG	2940
60	CGACCGATTA AAAATGCTGC TTATGCCACA CAGCACGGCT GGCAGGGGCG TGAGGGGTTG	3000
	AAGCCGTTGT ACGAGATGCA GGGAGCGCTG ATTAACAAC TGGATGCGCA TAACGTTCGT	3060

- 9 -

	CATAACGCGC CACAGCCAGA TTTGCAGAGC AAACCTGGAAA CTCTGGATTT AGGCGAACAT	3120
	GGCGCAGAAT TGCTTAACGA CATGAAGCGC TTCCGCGACG AACTGGAGCA GAGTGCAACC	3180
5	CGTTCGGTGA CCGTTTTAGG TCAACATCAG GGAGTGCTAA AAAGCAACGG TGAAATCAAT	3240
	AGCGAATTTA AGCCATCGCC CGGCAAGGCG TTGGTCCAGA GCTTTAACGT CAATCGCTCT	3300
10	GGTCAGGATC TAAGCAAGTC ACTGCAACAG GCAGTACATG CCACGCCGCC ATCCGCAGAG	3360
	AGTAAACTGC AATCCATGCT GGGGCACTTT GTCAGTGCCG GGGTGGATAT GAGTCATCAG	3420
	AAGGGCGAGA TCCCGCTGGG CCGCCAGCGC GATCCGAATG ATAAAACCGC ACTGACCAAA	3480
15	TCGCGTTTAA TTTTAGATAC CGTGACCATC GGTGAACTGC ATGAACTGGC CGATAAGGCG	3540
	AAACTGGTAT CTGACCATAA ACCCGATGCC GATCAGATAA AACAGCTGCG CCAGCAGTTC	3600
20	GATACGCTGC GTGAAAAGCG GTATGAGAGC AATCCGGTGA AGCATTACAC CGATATGGGC	3660
	TTCACCCATA ATAAGGCGCT GGAAGCAAAC TATGATGCGG TCAAAGCCTT TATCAATGCC	3720
	TTTAAGAAAG AGCACCACGG CGTCAATCTG ACCACGCGTA CCGTACTGGA ATCACAGGGC	3780
25	AGTGCGGAGC TGGCGAAGAA GCTCAAGAAT ACGCTGTTGT CCCTGGACAG TGGTGAAAGT	3840
	ATGAGCTTCA GCCGGTCATA TGGCGGGGGC GTCAGCACTG TCTTTGTGCC TACCCCTAGC	3900
30	AAGAAGGTGC CAGTTCGGT GATCCCCGGA GCCGGCATCA CGCTGGATCG CGCCTATAAC	3960
	CTGAGCTTCA GTCGTACCAG CGGCGGATTG AACGTCAGTT TTGGCCGCGA CGGCGGGGTG	4020
	AGTGGAACA TCATGGTCG TACCGGCCAT GATGTGATGC CCTATATGAC CGGTAAGAA	4080
35	ACCAGTGACG GTAACGCCAG TGAAGGTTG AGCGCAAAAC ATAAAATCAG CCCGGACTTG	4140
	CGTATCGGCG CTGCTGTGAG TGGCACCTG CAAGGAACGC TACAAAACAG CCTGAAGTTT	4200
40	AAGCTGACAG AGGATGAGCT GCCTGGCTTT ATCCATGGCT TGACGCATGG CACGTTGACC	4260
	CCGGCAGAAC TGTTGCAAAA GGGGATCGAA CATCAGATGA AGCAGGGCAG CAAACTGACG	4320
	TTTAGCGTCG ATACCTCGGC AAATCTGGAT CTGCGTGCCG GTATCAATCT GAACGAAGAC	4380
45	GGCAGTAAAC CAAATGGTGT CACTGCCCGT GTTCTGCGG GGCTAAGTGC ATCGGCAAAC	4440
	CTGGCCGCG GCTCGGTGA ACGCAGCACC ACCTCTGGCC AGTTTGGCAG CACGACTTCG	4500
50	GCCAGCAATA ACCGCCAAC CTTCTCAAC GGGGTCGGCG CGGGTGCTAA CCTGACGGCT	4560
	GCTTTAGGGG TTGCCATTG ATCTACGCAT GAAGGAAAC CGGTCGGGAT CTTCCCGGCA	4620
	TTTACCTCGA CCAATGTTTC GGCAGCGCTG GCGCTGGATA ACCGTACCTC ACAGAGTATC	4680
55	AGCCTGGAAT TGAAGCGCG GGAGCCGGTG ACCAGCAACG ATATCAGCGA GTTGACCTCC	4740
	ACGCTGGGAA AACACTTTAA GGATAGCGCC ACAACGAAGA TGCTTGCCGC TCTCAAAGAG	4800
60	TTAGATGACG CTAAGCCCGC TGAACAACTG CATATTTTAC AGCAGCATTT CAGTGCAAAA	4860
	GATGTCGTCG GTGATGAACG CTACGAGGCG GTGCGCAACC TGAAAAAACT GGTGATACGT	4920

- 10 -

CAACAGGCTG CGGACAGCCA CAGCATGGAA TTAGGATCTG CCAGTCACAG CACGACCTAC 4980
 AATAATCTGT CGAGAATAAA TAATGACGGC ATTGTCGAGC TGCTACACAA ACATTTTCGAT 5040
 5 GCGGCATTAC CAGCAAGCAG TGCCAAACGT CTGGGTGAAA TGATGAATAA CGATCCGGCA 5100
 CTGAAAGATA TTATTAAGCA GCTGCAAAGT ACGCCGTTCA GCAGCGCCAG CGTGTCGATG 5160
 10 GAGCTGAAAG ATGGTCTGCG TGAGCAGACG GAAAAAGCAA TACTGGACGG TAAGGTCGGT 5220
 CGTGAAGAAG TGGGAGTACT TTTCCAGGAT CGTAACAACT TGCCTGTTAA ATCGGTCAGC 5280
 GTCAGTCAGT CCGTCAGCAA AAGCGAAGGC TTCAATACCC CAGCGCTGTT ACTGGGGACG 5340
 15 AGCAACAGCG CTGCTATGAG CATGGAGCGC AACATCGGAA CCATTAATTT TAAATACGGC 5400
 CAGGATCAGA ACACCCACG GCGATTACC CTGGAGGGTG GAATAGCTCA GGCTAATCCG 5460
 CAGGTCGCAT CTGCGCTTAC TGATTTGAAG AAGGAAGGGC TGGAAATGAA GAGCTAA 5517
 20

This DNA molecule is known as the dspE gene. This isolated DNA molecule of the present invention encodes a protein or polypeptide which elicits a plant pathogen's hypersensitive response having an amino acid sequence of SEQ. ID. No. 2 as follows:

25 Met Glu Leu Lys Ser Leu Gly Thr Glu His Lys Ala Ala Val His Thr
 1 5 10 15
 30 Ala Ala His Asn Pro Val Gly His Gly Val Ala Leu Gln Gln Gly Ser
 20 25 30
 Ser Ser Ser Ser Pro Gln Asn Ala Ala Ala Ser Leu Ala Ala Glu Gly
 35 35 40 45
 35 Lys Asn Arg Gly Lys Met Pro Arg Ile His Gln Pro Ser Thr Ala Ala
 50 55 60
 40 Asp Gly Ile Ser Ala Ala His Gln Gln Lys Lys Ser Phe Ser Leu Arg
 65 70 75 80
 40 Gly Cys Leu Gly Thr Lys Lys Phe Ser Arg Ser Ala Pro Gln Gly Gln
 85 90 95
 45 Pro Gly Thr Thr His Ser Lys Gly Ala Thr Leu Arg Asp Leu Leu Ala
 100 105 110
 Arg Asp Asp Gly Glu Thr Gln His Glu Ala Ala Ala Pro Asp Ala Ala
 115 120 125
 50 Arg Leu Thr Arg Ser Gly Gly Val Lys Arg Arg Asn Met Asp Asp Met
 130 135 140
 55 Ala Gly Arg Pro Met Val Lys Gly Gly Ser Gly Glu Asp Lys Val Pro
 145 150 155 160
 Thr Gln Gln Lys Arg His Gln Leu Asn Asn Phe Gly Gln Met Arg Gln
 165 170 175

- 11 -

	Thr Met Leu Ser Lys Met Ala His Pro Ala Ser Ala Asn Ala Gly Asp	
	180	185 190
5	Arg Leu Gln His Ser Pro Pro His Ile Pro Gly Ser His His Glu Ile	
	195	200 205
	Lys Glu Glu Pro Val Gly Ser Thr Ser Lys Ala Thr Thr Ala His Ala	
	210	215 220
10	Asp Arg Val Glu Ile Ala Gln Glu Asp Asp Asp Ser Glu Phe Gln Gln	
	225	230 235 240
	Leu His Gln Gln Arg Leu Ala Arg Glu Arg Glu Asn Pro Pro Gln Pro	
	245	250 255
15	Pro Lys Leu Gly Val Ala Thr Pro Ile Ser Ala Arg Phe Gln Pro Lys	
	260	265 270
20	Leu Thr Ala Val Ala Glu Ser Val Leu Glu Gly Thr Asp Thr Thr Gln	
	275	280 285
	Ser Pro Leu Lys Pro Gln Ser Met Leu Lys Gly Ser Gly Ala Gly Val	
	290	295 300
25	Thr Pro Leu Ala Val Thr Leu Asp Lys Gly Lys Leu Gln Leu Ala Pro	
	305	310 315 320
	Asp Asn Pro Pro Ala Leu Asn Thr Leu Leu Lys Gln Thr Leu Gly Lys	
	325	330 335
30	Asp Thr Gln His Tyr Leu Ala His His Ala Ser Ser Asp Gly Ser Gln	
	340	345 350
35	His Leu Leu Leu Asp Asn Lys Gly His Leu Phe Asp Ile Lys Ser Thr	
	355	360 365
	Ala Thr Ser Tyr Ser Val Leu His Asn Ser His Pro Gly Glu Ile Lys	
	370	375 380
40	Gly Lys Leu Ala Gln Ala Gly Thr Gly Ser Val Ser Val Asp Gly Lys	
	385	390 395 400
	Ser Gly Lys Ile Ser Leu Gly Ser Gly Thr Gln Ser His Asn Lys Thr	
	405	410 415
45	Met Leu Ser Gln Pro Gly Glu Ala His Arg Ser Leu Leu Thr Gly Ile	
	420	425 430
50	Trp Gln His Pro Ala Gly Ala Ala Arg Pro Gln Gly Glu Ser Ile Arg	
	435	440 445
	Leu His Asp Asp Lys Ile His Ile Leu His Pro Glu Leu Gly Val Trp	
	450	455 460
55	Gln Ser Ala Asp Lys Asp Thr His Ser Gln Leu Ser Arg Gln Ala Asp	
	465	470 475 480
	Gly Lys Leu Tyr Ala Leu Lys Asp Asn Arg Thr Leu Gln Asn Leu Ser	
	485	490 495
60	Asp Asn Lys Ser Ser Glu Lys Leu Val Asp Lys Ile Lys Ser Tyr Ser	
	500	505 510

- 12 -

	Val Asp Gln Arg Gly Gln Val Ala Ile Leu Thr Asp Thr Pro Gly Arg	
	515	520 525
5	His Lys Met Ser Ile Met Pro Ser Leu Asp Ala Ser Pro Glu Ser His	
	530	535 540
	Ile Ser Leu Ser Leu His Phe Ala Asp Ala His Gln Gly Leu Leu His	
10	545	550 555 560
	Gly Lys Ser Glu Leu Glu Ala Gln Ser Val Ala Ile Ser His Gly Arg	
	565	570 575
15	Leu Val Val Ala Asp Ser Glu Gly Lys Leu Phe Ser Ala Ala Ile Pro	
	580	585 590
	Lys Gln Gly Asp Gly Asn Glu Leu Lys Met Lys Ala Met Pro Gln His	
	595	600 605
20	Ala Leu Asp Glu His Phe Gly His Asp His Gln Ile Ser Gly Phe Phe	
	610	615 620
	His Asp Asp His Gly Gln Leu Asn Ala Leu Val Lys Asn Asn Phe Arg	
25	625	630 635 640
	Gln Gln His Ala Cys Pro Leu Gly Asn Asp His Gln Phe His Pro Gly	
	645	650 655
30	Trp Asn Leu Thr Asp Ala Leu Val Ile Asp Asn Gln Leu Gly Leu His	
	660	665 670
	His Thr Asn Pro Glu Pro His Glu Ile Leu Asp Met Gly His Leu Gly	
	675	680 685
35	Ser Leu Ala Leu Gln Glu Gly Lys Leu His Tyr Phe Asp Gln Leu Thr	
	690	695 700
	Lys Gly Trp Thr Gly Ala Glu Ser Asp Cys Lys Gln Leu Lys Lys Gly	
40	705	710 715 720
	Leu Asp Gly Ala Ala Tyr Leu Leu Lys Asp Gly Glu Val Lys Arg Leu	
	725	730 735
45	Asn Ile Asn Gln Ser Thr Ser Ser Ile Lys His Gly Thr Glu Asn Val	
	740	745 750
	Phe Ser Leu Pro His Val Arg Asn Lys Pro Glu Pro Gly Asp Ala Leu	
	755	760 765
50	Gln Gly Leu Asn Lys Asp Asp Lys Ala Gln Ala Met Ala Val Ile Gly	
	770	775 780
	Val Asn Lys Tyr Leu Ala Leu Thr Glu Lys Gly Asp Ile Arg Ser Phe	
55	785	790 795 800
	Gln Ile Lys Pro Gly Thr Gln Gln Leu Glu Arg Pro Ala Gln Thr Leu	
	805	810 815
60	Ser Arg Glu Gly Ile Ser Gly Glu Leu Lys Asp Ile His Val Asp His	
	820	825 830

- 13 -

Lys Gln Asn Leu Tyr Ala Leu Thr His Glu Gly Glu Val Phe His Gln
 835 840 845
 5 Pro Arg Glu Ala Trp Gln Asn Gly Ala Glu Ser Ser Ser Trp His Lys
 850 855 860
 Leu Ala Leu Pro Gln Ser Glu Ser Lys Leu Lys Ser Leu Asp Met Ser
 865 870 875 880
 10 His Glu His Lys Pro Ile Ala Thr Phe Glu Asp Gly Ser Gln His Gln
 885 890 895
 Leu Lys Ala Gly Gly Trp His Ala Tyr Ala Ala Pro Glu Arg Gly Pro
 900 905 910
 15 Leu Ala Val Gly Thr Ser Gly Ser Gln Thr Val Phe Asn Arg Leu Met
 915 920 925
 Gln Gly Val Lys Gly Lys Val Ile Pro Gly Ser Gly Leu Thr Val Lys
 20 930 935 940
 Leu Ser Ala Gln Thr Gly Gly Met Thr Gly Ala Glu Gly Arg Lys Val
 945 950 955 960
 25 Ser Ser Lys Phe Ser Glu Arg Ile Arg Ala Tyr Ala Phe Asn Pro Thr
 965 970 975
 Met Ser Thr Pro Arg Pro Ile Lys Asn Ala Ala Tyr Ala Thr Gln His
 30 980 985 990
 Gly Trp Gln Gly Arg Glu Gly Leu Lys Pro Leu Tyr Glu Met Gln Gly
 995 1000 1005
 35 Ala Leu Ile Lys Gln Leu Asp Ala His Asn Val Arg His Asn Ala Pro
 1010 1015 1020
 Gln Pro Asp Leu Gln Ser Lys Leu Glu Thr Leu Asp Leu Gly Glu His
 1025 1030 1035 1040
 40 Gly Ala Glu Leu Leu Asn Asp Met Lys Arg Phe Arg Asp Glu Leu Glu
 1045 1050 1055
 Gln Ser Ala Thr Arg Ser Val Thr Val Leu Gly Gln His Gln Gly Val
 45 1060 1065 1070
 Leu Lys Ser Asn Gly Glu Ile Asn Ser Glu Phe Lys Pro Ser Pro Gly
 1075 1080 1085
 50 Lys Ala Leu Val Gln Ser Phe Asn Val Asn Arg Ser Gly Gln Asp Leu
 1090 1095 1100
 Ser Lys Ser Leu Gln Gln Ala Val His Ala Thr Pro Pro Ser Ala Glu
 1105 1110 1115 1120
 55 Ser Lys Leu Gln Ser Met Leu Gly His Phe Val Ser Ala Gly Val Asp
 1125 1130 1135
 Met Ser His Gln Lys Gly Glu Ile Pro Leu Gly Arg Gln Arg Asp Pro
 1140 1145 1150
 60 Asn Asp Lys Thr Ala Leu Thr Lys Ser Arg Leu Ile Leu Asp Thr Val
 1155 1160 1165

- 14 -

Thr Ile Gly Glu Leu His Glu Leu Ala Asp Lys Ala Lys Leu Val Ser
 1170 1175 1180

5 Asp His Lys Pro Asp Ala Asp Gln Ile Lys Gln Leu Arg Gln Gln Phe
 1185 1190 1195 1200

Asp Thr Leu Arg Glu Lys Arg Tyr Glu Ser Asn Pro Val Lys His Tyr
 1205 1210 1215

10 Thr Asp Met Gly Phe Thr His Asn Lys Ala Leu Glu Ala Asn Tyr Asp
 1220 1225 1230

Ala Val Lys Ala Phe Ile Asn Ala Phe Lys Lys Glu His His Gly Val
 1235 1240 1245

15 Asn Leu Thr Thr Arg Thr Val Leu Glu Ser Gln Gly Ser Ala Glu Leu
 1250 1255 1260

20 Ala Lys Lys Leu Lys Asn Thr Leu Leu Ser Leu Asp Ser Gly Glu Ser
 1265 1270 1275 1280

Met Ser Phe Ser Arg Ser Tyr Gly Gly Gly Val Ser Thr Val Phe Val
 1285 1290 1295

25 Pro Thr Leu Ser Lys Lys Val Pro Val Pro Val Ile Pro Gly Ala Gly
 1300 1305 1310

30 Ile Thr Leu Asp Arg Ala Tyr Asn Leu Ser Phe Ser Arg Thr Ser Gly
 1315 1320 1325

Gly Leu Asn Val Ser Phe Gly Arg Asp Gly Gly Val Ser Gly Asn Ile
 1330 1335 1340

35 Met Val Ala Thr Gly His Asp Val Met Pro Tyr Met Thr Gly Lys Lys
 1345 1350 1355 1360

Thr Ser Ala Gly Asn Ala Ser Asp Trp Leu Ser Ala Lys His Lys Ile
 1365 1370 1375

40 Ser Pro Asp Leu Arg Ile Gly Ala Ala Val Ser Gly Thr Leu Gln Gly
 1380 1385 1390

Thr Leu Gln Asn Ser Leu Lys Phe Lys Leu Thr Glu Asp Glu Leu Pro
 1395 1400 1405

Gly Phe Ile His Gly Leu Thr His Gly Thr Leu Thr Pro Ala Glu Leu
 1410 1415 1420

50 Leu Gln Lys Gly Ile Glu His Gln Met Lys Gln Gly Ser Lys Leu Thr
 1425 1430 1435 1440

Phe Ser Val Asp Thr Ser Ala Asn Leu Asp Leu Arg Ala Gly Ile Asn
 1445 1450 1455

55 Leu Asn Glu Asp Gly Ser Lys Pro Asn Gly Val Thr Ala Arg Val Ser
 1460 1465 1470

60 Ala Gly Leu Ser Ala Ser Ala Asn Leu Ala Ala Gly Ser Arg Glu Arg
 1475 1480 1485

- 15 -

	Ser Thr Thr Ser Gly Gln Phe Gly Ser Thr Thr Ser Ala Ser Asn Asn	
	1490	1495 1500
5	Arg Pro Thr Phe Leu Asn Gly Val Gly Ala Gly Ala Asn Leu Thr Ala	
	1505	1510 1515 1520
	Ala Leu Gly Val Ala His Ser Ser Thr His Glu Gly Lys Pro Val Gly	
		1525 1530 1535
10	Ile Phe Pro Ala Phe Thr Ser Thr Asn Val Ser Ala Ala Leu Ala Leu	
		1540 1545 1550
	Asp Asn Arg Thr Ser Gln Ser Ile Ser Leu Glu Leu Lys Arg Ala Glu	
15		1555 1560 1565
	Pro Val Thr Ser Asn Asp Ile Ser Glu Leu Thr Ser Thr Leu Gly Lys	
		1570 1575 1580
20	His Phe Lys Asp Ser Ala Thr Thr Lys Met Leu Ala Ala Leu Lys Glu	
		1585 1590 1595 1600
	Leu Asp Asp Ala Lys Pro Ala Glu Gln Leu His Ile Leu Gln Gln His	
		1605 1610 1615
25	Phe Ser Ala Lys Asp Val Val Gly Asp Glu Arg Tyr Glu Ala Val Arg	
		1620 1625 1630
	Asn Leu Lys Lys Leu Val Ile Arg Gln Gln Ala Ala Asp Ser His Ser	
30		1635 1640 1645
	Met Glu Leu Gly Ser Ala Ser His Ser Thr Thr Tyr Asn Asn Leu Ser	
		1650 1655 1660
35	Arg Ile Asn Asn Asp Gly Ile Val Glu Leu Leu His Lys His Phe Asp	
		1665 1670 1675 1680
	Ala Ala Leu Pro Ala Ser Ser Ala Lys Arg Leu Gly Glu Met Met Asn	
		1685 1690 1695
40	Asn Asp Pro Ala Leu Lys Asp Ile Ile Lys Gln Leu Gln Ser Thr Pro	
		1700 1705 1710
	Phe Ser Ser Ala Ser Val Ser Met Glu Leu Lys Asp Gly Leu Arg Glu	
45		1715 1720 1725
	Gln Thr Glu Lys Ala Ile Leu Asp Gly Lys Val Gly Arg Glu Glu Val	
		1730 1735 1740
50	Gly Val Leu Phe Gln Asp Arg Asn Asn Leu Arg Val Lys Ser Val Ser	
		1745 1750 1755 1760
	Val Ser Gln Ser Val Ser Lys Ser Glu Gly Phe Asn Thr Pro Ala Leu	
		1765 1770 1775
55	Leu Leu Gly Thr Ser Asn Ser Ala Ala Met Ser Met Glu Arg Asn Ile	
		1780 1785 1790
	Gly Thr Ile Asn Phe Lys Tyr Gly Gln Asp Gln Asn Thr Pro Arg Arg	
60		1795 1800 1805
	Phe Thr Leu Glu Gly Gly Ile Ala Gln Ala Asn Pro Gln Val Ala Ser	
		1810 1815 1820

- 16 -

Ala Leu Thr Asp Leu Lys Lys Glu Gly Leu Glu Met Lys Ser
 1825 1830 1835

5

This protein or polypeptide is about 198 kDa and has a pI of 8.98.

The present invention relates to an isolated DNA molecule having a nucleotide sequence of SEQ. ID. No. 3 as follows:

10

ATGACATCGT CACAGCAGCG GGTGAAAGG TTTTACAGT ATTTCTCCGC CGGGTGTAAG 60
 ACGCCCATAC ATCTGAAAGA CGGGGTGTGC GCCCTGTATA ACGAACAAGA TGAGGAGGCG 120
 15 GCGGTGCTGG AAGTACCGCA ACACAGCGAC AGCCTGTAC TACACTGCCG AATCATTGAG 180
 GCTGACCCAC AAACCTCAAT AACCTGTAT TCGATGCTAT TACAGCTGAA TTTTGAAATG 240
 GCGGCCATGC GCGGCTGTTG GCTGGCGCTG GATGAACTGC ACAACGTGCG TTTATGTTTT 300
 20 CAGCAGTCGC TGGAGCATCT GGATGAAGCA AGTTTAGCG ATATCGTTAG CGGCITCATC 360
 GAACATGCGG CAGAAGTGCG TGAGTATATA GCGCAATTAG ACGAGAGTAG CGCGGCATAA 420

25

This is known as the dspF gene. This isolated DNA molecule of the present invention encodes a hypersensitive response elicitor protein or polypeptide having an amino acid sequence of SEQ. ID. No. 4 as follows:

30

Met Thr Ser Ser Gln Gln Arg Val Glu Arg Phe Leu Gln Tyr Phe Ser
 1 5 10 15
 Ala Gly Cys Lys Thr Pro Ile His Leu Lys Asp Gly Val Cys Ala Leu
 20 25 30
 35 Tyr Asn Glu Gln Asp Glu Glu Ala Ala Val Leu Glu Val Pro Gln His
 35 40 45
 40 Ser Asp Ser Leu Leu Leu His Cys Arg Ile Ile Glu Ala Asp Pro Gln
 50 55 60
 Thr Ser Ile Thr Leu Tyr Ser Met Leu Leu Gln Leu Asn Phe Glu Met
 65 70 75 80
 45 Ala Ala Met Arg Gly Cys Trp Leu Ala Leu Asp Glu Leu His Asn Val
 85 90 95
 Arg Leu Cys Phe Gln Gln Ser Leu Glu His Leu Asp Glu Ala Ser Phe
 100 105 110
 50 Ser Asp Ile Val Ser Gly Phe Ile Glu His Ala Ala Glu Val Arg Glu
 115 120 125
 Tyr Ile Ala Gln Leu Asp Glu Ser Ser Ala Ala
 130 135

55

- 17 -

This protein or polypeptide is about 16 kDa and has a pI of 4.45.

Fragments of the above hypersensitive response elicitor polypeptide or protein are encompassed by the present invention.

5 Suitable fragments can be produced by several means. In the first, subclones of the gene encoding the elicitor protein of the present invention are produced by conventional molecular genetic manipulation by subcloning gene fragments. The subclones then are expressed *in vitro* or *in vivo* in bacterial cells to yield a smaller protein or peptide that can be tested for elicitor activity according to
10 the procedure described below.

As an alternative, fragments of an elicitor protein can be produced by digestion of a full-length elicitor protein with proteolytic enzymes like chymotrypsin or *Staphylococcus* proteinase A, or trypsin. Different proteolytic enzymes are likely to cleave elicitor proteins at different sites based on the amino acid sequence of the
15 elicitor protein. Some of the fragments that result from proteolysis may be active elicitors of resistance.

In another approach, based on knowledge of the primary structure of the protein, fragments of the elicitor protein gene may be synthesized by using the PCR technique together with specific sets of primers chosen to represent particular
20 portions of the protein. These then would be cloned into an appropriate vector for increased expression of a truncated peptide or protein.

Chemical synthesis can also be used to make suitable fragments. Such a synthesis is carried out using known amino acid sequences for the elicitor being produced. Alternatively, subjecting a full length elicitor to high temperatures and
25 pressures will produce fragments. These fragments can then be separated by conventional procedures (e.g., chromatography, SDS-PAGE).

Variants may also (or alternatively) be modified by, for example, the deletion or addition of amino acids that have minimal influence on the properties, secondary structure and hydrophobic nature of the polypeptide. For example, a
30 polypeptide may be conjugated to a signal (or leader) sequence at the N-terminal end of the protein which co-translationally or post-translationally directs transfer of the protein. The polypeptide may also be conjugated to a linker or other sequence for ease of synthesis, purification, or identification of the polypeptide.

- 18 -

Suitable DNA molecules are those that hybridize to a DNA molecule comprising a nucleotide sequence of SEQ. ID. Nos. 1 and 3, under stringent conditions. An example of suitable high stringency conditions is when hybridization is carried out at 65°C for 20 hours in a medium containing 1M NaCl, 50 mM Tris-HCl, pH 7.4, 10 mM EDTA, 0.1% sodium dodecyl sulfate, 0.2% ficoll, 0.2%
5 polyvinylpyrrolidone, 0.2% bovine serum albumin, 50 µm g/ml *E. coli* DNA. However, any DNA molecules hybridizing to a DNA molecule comprising the nucleotide sequences of SEQ. ID. Nos. 1 and 3, under such stringent conditions must not be identical to the nucleic acids encoding the hypersensitive response elicitor
10 proteins or polypeptides of *E. amylovora* (as disclosed by Wei, Z.-M., et al., "Harpin, Elicitor of the Hypersensitive Response Produced by the Plant Pathogen *Erwinia amylovora*," Science 257:85-88 (1992), which is hereby incorporated by reference), *Erwinia chrysanthemi* (as disclosed by Bauer, et. al., "*Erwinia chrysanthemi* Harpin_{Ech}: Soft-Rot Pathogenesis," MPMI 8(4): 484-91 (1995), which is hereby
15 incorporated by reference), *Erwinia carotovora* (as disclosed by Cui, et. al., "The RsmA⁻ Mutants of *Erwinia carotovora* subsp. *carotovora* Strain Ecc71 Overexpress *hrpN*_{Ecc} and Elicit a Hypersensitive Reaction-like Response in Tobacco Leaves," MPMI 9(7): 565-73 (1966), which is hereby incorporated by reference), *Erwinia stewartii* (as disclosed by Ahmad, et. al., "Harpin is not Necessary for the
20 Pathogenicity of *Erwinia stewartii* on Maize," 8th Int'l. Cong. Molec. Plant-Microb. Inter. July 14-19, 1996 and Ahmad, et. al., "Harpin is not Necessary for the Pathogenicity of *Erwinia stewartii* on Maize," Ann. Mtg. Am. Phytopath. Soc. July 27-31, 1996), which are hereby incorporated by reference), and *Pseudomonas syringae* pv. *syringae* (WO 94/26782 to Cornell Research Foundation, Inc., which is
25 hereby incorporated by reference).

The protein or polypeptide of the present invention is preferably produced in purified form (preferably at least about 80%, more preferably 90%, pure) by conventional techniques. Typically, the protein or polypeptide of the present invention is secreted into the growth medium of recombinant host cells.
30 Alternatively, the protein or polypeptide of the present invention is produced but not secreted into growth medium. In such cases, to isolate the protein, the host cell (e.g., *E. coli*) carrying a recombinant plasmid is propagated, lysed by sonication, heat, or

- 19 -

chemical treatment, and the homogenate is centrifuged to remove bacterial debris.

The supernatant is then subjected to sequential ammonium sulfate precipitation. The fraction containing the polypeptide or protein of the present invention is subjected to gel filtration in an appropriately sized dextran or polyacrylamide column to separate the proteins. If necessary, the protein fraction may be further purified by HPLC.

The DNA molecule encoding the hypersensitive response elicitor polypeptide or protein can be incorporated in cells using conventional recombinant DNA technology. Generally, this involves inserting the DNA molecule into an expression system to which the DNA molecule is heterologous (i.e. not normally present). The heterologous DNA molecule is inserted into the expression system or vector in proper sense orientation and correct reading frame. The vector contains the necessary elements for the transcription and translation of the inserted protein-coding sequences.

U.S. Patent No. 4,237,224 to Cohen and Boyer, which is hereby incorporated by reference, describes the production of expression systems in the form of recombinant plasmids using restriction enzyme cleavage and ligation with DNA ligase. These recombinant plasmids are then introduced by means of transformation and replicated in unicellular cultures including procaryotic organisms and eucaryotic cells grown in tissue culture.

Recombinant genes may also be introduced into viruses, such as vaccinia virus. Recombinant viruses can be generated by transfection of plasmids into cells infected with virus.

Suitable vectors include, but are not limited to, the following viral vectors such as lambda vector system gt11, gt WES.tB, Charon 4, and plasmid vectors such as pBR322, pBR325, pACYC177, pACYC1084, pUC8, pUC9, pUC18, pUC19, pLG339, pR290, pKC37, pKC101, SV 40, pBluescript II SK +/- or KS +/- (see "Stratagene Cloning Systems" Catalog (1993) from Stratagene, La Jolla, Calif, which is hereby incorporated by reference), pQE, pIH821, pGEX, pET series (see F.W. Studier et. al., "Use of T7 RNA Polymerase to Direct Expression of Cloned Genes," Gene Expression Technology vol. 185 (1990), which is hereby incorporated by reference), and any derivatives thereof. Recombinant molecules can be introduced into cells via transformation, particularly transduction, conjugation, mobilization, or

- 20 -

electroporation. The DNA sequences are cloned into the vector using standard cloning procedures in the art, as described by Sambrook et al., Molecular Cloning: A Laboratory Manual, Cold Springs Laboratory, Cold Springs Harbor, New York (1989), which is hereby incorporated by reference.

5 A variety of host-vector systems may be utilized to express the protein-encoding sequence(s). Primarily, the vector system must be compatible with the host cell used. Host-vector systems include but are not limited to the following: bacteria transformed with bacteriophage DNA, plasmid DNA, or cosmid DNA; microorganisms such as yeast containing yeast vectors; mammalian cell systems
10 infected with virus (e.g., vaccinia virus, adenovirus, etc.); insect cell systems infected with virus (e.g., baculovirus); and plant cells infected by bacteria. The expression elements of these vectors vary in their strength and specificities. Depending upon the host-vector system utilized, any one of a number of suitable transcription and translation elements can be used.

15 Different genetic signals and processing events control many levels of gene expression (e.g., DNA transcription and messenger RNA (mRNA) translation).

 Transcription of DNA is dependent upon the presence of a promotor which is a DNA sequence that directs the binding of RNA polymerase and thereby promotes mRNA synthesis. The DNA sequences of eucaryotic promoters differ from
20 those of procaryotic promoters. Furthermore, eucaryotic promoters and accompanying genetic signals may not be recognized in or may not function in a procaryotic system, and, further, procaryotic promoters are not recognized and do not function in eucaryotic cells.

 Similarly, translation of mRNA in procaryotes depends upon
25 the presence of the proper procaryotic signals which differ from those of eucaryotes. Efficient translation of mRNA in procaryotes requires a ribosome binding site called the Shine-Dalgarno ("SD") sequence on the mRNA. This sequence is a short nucleotide sequence of mRNA that is located before the start codon, usually AUG, which encodes the amino-terminal methionine of the protein. The SD sequences are
30 complementary to the 3'-end of the 16S rRNA (ribosomal RNA) and probably promote binding of mRNA to ribosomes by duplexing with the rRNA to allow correct positioning of the ribosome. For a review on maximizing gene expression, see

- 21 -

Roberts and Lauer, Methods in Enzymology, 68:473 (1979), which is hereby incorporated by reference.

Promoters vary in their "strength" (i.e. their ability to promote transcription). For the purposes of expressing a cloned gene, it is desirable to use strong promoters in order to obtain a high level of transcription and, hence, expression of the gene. Depending upon the host cell system utilized, any one of a number of suitable promoters may be used. For instance, when cloning in *E. coli*, its bacteriophages, or plasmids, promoters such as the T7 phage promoter, *lac* promoter, *trp* promoter, *recA* promoter, ribosomal RNA promoter, the P_R and P_L promoters of coliphage lambda and others, including but not limited, to *lacUV5*, *ompF*, *bla*, *lpp*, and the like, may be used to direct high levels of transcription of adjacent DNA segments. Additionally, a hybrid *trp-lacUV5 (tac)* promoter or other *E. coli* promoters produced by recombinant DNA or other synthetic DNA techniques may be used to provide for transcription of the inserted gene.

Bacterial host cell strains and expression vectors may be chosen which inhibit the action of the promoter unless specifically induced. In certain operations, the addition of specific inducers is necessary for efficient transcription of the inserted DNA. For example, the *lac* operon is induced by the addition of lactose or IPTG (isopropylthio-beta-D-galactoside). A variety of other operons, such as *trp*, *pro*, etc., are under different controls.

Specific initiation signals are also required for efficient gene transcription and translation in procaryotic cells. These transcription and translation initiation signals may vary in "strength" as measured by the quantity of gene specific messenger RNA and protein synthesized, respectively. The DNA expression vector, which contains a promoter, may also contain any combination of various "strong" transcription and/or translation initiation signals. For instance, efficient translation in *E. coli* requires an SD sequence about 7-9 bases 5' to the initiation codon ("ATG") to provide a ribosome binding site. Thus, any SD-ATG combination that can be utilized by host cell ribosomes may be employed. Such combinations include but are not limited to the SD-ATG combination from the *cro* gene or the *N* gene of coliphage lambda, or from the *E. coli* tryptophan E, D, C, B or A genes. Additionally, any SD-

- 22 -

ATG combination produced by recombinant DNA or other techniques involving incorporation of synthetic nucleotides may be used.

Once the isolated DNA molecule encoding the hypersensitive response elicitor polypeptide or protein has been cloned into an expression system, it is ready
5 to be incorporated into a host cell. Such incorporation can be carried out by the various forms of transformation noted above, depending upon the vector/host cell system. Suitable host cells include, but are not limited to, bacteria, virus, yeast, mammalian cells, insect, plant, and the like.

The present invention further relates to methods of imparting disease
10 resistance to plants, enhancing plant growth, and/or effecting insect control for plants. These methods involve applying a hypersensitive response elicitor polypeptide or protein in a non-infectious form to all or part of a plant or a plant seed under conditions where the polypeptide or protein contacts all or part of the cells of the plant or plant seed. Alternatively, the hypersensitive response elicitor protein or
15 polypeptide can be applied to plants such that seeds recovered from such plants themselves are able to impart disease resistance in plants, to enhance plant growth, and/or to effect insect control.

As an alternative to applying a hypersensitive response elicitor polypeptide or protein to plants or plant seeds in order to impart disease resistance in
20 plants, to effect plant growth, and/or to control insects on the plants or plants grown from the seeds, transgenic plants or plant seeds can be utilized. When utilizing transgenic plants, this involves providing a transgenic plant transformed with a DNA molecule encoding a hypersensitive response elicitor polypeptide or protein and growing the plant under conditions effective to permit that DNA molecule to impart
25 disease resistance to plants, to enhance plant growth, and/or to control insects. Alternatively, a transgenic plant seed transformed with a DNA molecule encoding a hypersensitive response elicitor polypeptide or protein can be provided and planted in soil. A plant is then propagated from the planted seed under conditions effective to permit that DNA molecule to impart disease resistance to plants, to enhance plant
30 growth, and/or to control insects.

The embodiment of the present invention where the hypersensitive response elicitor polypeptide or protein is applied to the plant or plant seed can be

- 23 -

carried out in a number of ways, including: 1) application of an isolated elicitor polypeptide or protein; 2) application of bacteria which do not cause disease and are transformed with genes encoding a hypersensitive response elicitor polypeptide or protein; and 3) application of bacteria which cause disease in some plant species (but not in those to which they are applied) and naturally contain a gene encoding the hypersensitive response elicitor polypeptide or protein.

In one embodiment of the present invention, the hypersensitive response elicitor polypeptide or protein of the present invention can be isolated from *Erwinia amylovora* as described in the Examples *infra*. Preferably, however, the isolated hypersensitive response elicitor polypeptide or protein of the present invention is produced recombinantly and purified as described *supra*.

In other embodiments of the present invention, the hypersensitive response elicitor polypeptide or protein of the present invention can be applied to plants or plant seeds by applying bacteria containing genes encoding the hypersensitive response elicitor polypeptide or protein. Such bacteria must be capable of secreting or exporting the polypeptide or protein so that the elicitor can contact plant or plant seed cells. In these embodiments, the hypersensitive response elicitor polypeptide or protein is produced by the bacteria *in planta* or on seeds or just prior to introduction of the bacteria to the plants or plant seeds.

In one embodiment of the bacterial application mode of the present invention, the bacteria do not cause the disease and have been transformed (e.g., recombinantly) with genes encoding a hypersensitive response elicitor polypeptide or protein. For example, *E. coli*, which does not elicit a hypersensitive response in plants, can be transformed with genes encoding a hypersensitive response elicitor polypeptide or protein and then applied to plants. Bacterial species other than *E. coli* can also be used in this embodiment of the present invention.

In another embodiment of the bacterial application mode of the present invention, the bacteria do cause disease and naturally contain a gene encoding a hypersensitive response elicitor polypeptide or protein. Examples of such bacteria are noted above. However, in this embodiment, these bacteria are applied to plants or their seeds which are not susceptible to the disease carried by the bacteria. For example, *Erwinia amylovora* causes disease in apple or pear but not in tomato.

- 24 -

However, such bacteria will elicit a hypersensitive response in tomato. Accordingly, in accordance with this embodiment of the present invention, *Erwinia amylovora* can be applied to tomato plants or seeds to enhance growth without causing disease in that species.

5 The method of the present invention can be utilized to treat a wide variety of plants or their seeds to impart disease resistance, enhance growth, and/or control insects. Suitable plants include dicots and monocots. More particularly, useful crop plants can include: alfalfa, rice, wheat, barley, rye, cotton, sunflower, peanut, corn, potato, sweet potato, bean, pea, chicory, lettuce, endive, cabbage,
10 brussel sprout, beet, parsnip, turnip, cauliflower, broccoli, turnip, radish, spinach, onion, garlic, eggplant, pepper, celery, carrot, squash, pumpkin, zucchini, cucumber, apple, pear, melon, citrus, strawberry, grape, raspberry, pineapple, soybean, tobacco, tomato, sorghum, and sugarcane. Examples of suitable ornamental plants are:
15 *Arabidopsis thaliana*, *Saintpaulia*, petunia, pelargonium, poinsettia, chrysanthemum, carnation, and zinnia.

 With regard to the use of the hypersensitive response elicitor protein or polypeptide of the present invention in imparting disease resistance, absolute immunity against infection may not be conferred, but the severity of the disease is reduced and symptom development is delayed. Lesion number, lesion size, and
20 extent of sporulation of fungal pathogens are all decreased. This method of imparting disease resistance has the potential for treating previously untreatable diseases, treating diseases systemically which might not be treated separately due to cost, and avoiding the use of infectious agents or environmentally harmful materials.

 The method of imparting pathogen resistance to plants in accordance
25 with the present invention is useful in imparting resistance to a wide variety of pathogens including viruses, bacteria, and fungi. Resistance, *inter alia*, to the following viruses can be achieved by the method of the present invention: *Tobacco mosaic virus* and *Tomato mosaic virus*. Resistance, *inter alia*, to the following bacteria can also be imparted to plants in accordance with present invention:
30 *Pseudomonas solanacearum*, *Pseudomonas syringae* pv. *tabaci*, and *Xanthamonas campestris* pv. *pelargonii*. Plants can be made resistant, *inter alia*, to the following

- 25 -

fungi by use of the method of the present invention: *Fusarium oxysporum* and *Phytophthora infestans*.

With regard to the use of the hypersensitive response elicitor protein or polypeptide of the present invention to enhance plant growth, various forms of plant growth enhancement or promotion can be achieved. This can occur as early as when plant growth begins from seeds or later in the life of a plant. For example, plant growth according to the present invention encompasses greater yield, increased quantity of seeds produced, increased percentage of seeds germinated, increased plant size, greater biomass, more and bigger fruit, earlier fruit coloration, and earlier fruit and plant maturation. As a result, the present invention provides significant economic benefit to growers. For example, early germination and early maturation permit crops to be grown in areas where short growing seasons would otherwise preclude their growth in that locale. Increased percentage of seed germination results in improved crop stands and more efficient seed use. Greater yield, increased size, and enhanced biomass production allow greater revenue generation from a given plot of land.

Another aspect of the present invention is directed to effecting any form of insect control for plants. For example, insect control according to the present invention encompasses preventing insects from contacting plants to which the hypersensitive response elicitor has been applied, preventing direct insect damage to plants by feeding injury, causing insects to depart from such plants, killing insects proximate to such plants, interfering with insect larval feeding on such plants, preventing insects from colonizing host plants, preventing colonizing insects from releasing phytotoxins, etc. The present invention also prevents subsequent disease damage to plants resulting from insect infection.

The present invention is effective against a wide variety of insects. European corn borer is a major pest of corn (dent and sweet corn) but also feeds on over 200 plant species including green, wax, and lima beans and edible soybeans, peppers, potato, and tomato plus many weed species. Additional insect larval feeding pests which damage a wide variety of vegetable crops include the following: beet armyworm, cabbage looper, corn ear worm, fall armyworm, diamondback moth, cabbage root maggot, onion maggot, seed corn maggot, pickleworm (melonworm), pepper maggot, and tomato pinworm. Collectively, this group of insect pests

- 26 -

represents the most economically important group of pests for vegetable production worldwide.

The method of the present invention involving application of the hypersensitive response elicitor polypeptide or protein can be carried out through a variety of procedures when all or part of the plant is treated, including leaves, stems, roots, propagules (e.g., cuttings), etc. This may (but need not) involve infiltration of the hypersensitive response elicitor polypeptide or protein into the plant. Suitable application methods include high or low pressure spraying, injection, and leaf abrasion proximate to when elicitor application takes place. When treating plant seeds, in accordance with the application embodiment of the present invention, the hypersensitive response elicitor protein or polypeptide can be applied by low or high pressure spraying, coating, immersion, or injection. Other suitable application procedures can be envisioned by those skilled in the art provided they are able to effect contact of the hypersensitive response elicitor polypeptide or protein with cells of the plant or plant seed. Once treated with the hypersensitive response elicitor of the present invention, the seeds can be planted in natural or artificial soil and cultivated using conventional procedures to produce plants. After plants have been propagated from seeds treated in accordance with the present invention, the plants may be treated with one or more applications of the hypersensitive response elicitor protein or polypeptide to impart disease resistance to plants, to enhance plant growth, and/or to control insects on the plants.

The hypersensitive response elicitor polypeptide or protein can be applied to plants or plant seeds in accordance with the present invention alone or in a mixture with other materials. Alternatively, the hypersensitive response elicitor polypeptide or protein can be applied separately to plants with other materials being applied at different times.

A composition suitable for treating plants or plant seeds in accordance with the application embodiment of the present invention contains a hypersensitive response elicitor polypeptide or protein in a carrier. Suitable carriers include water, aqueous solutions, slurries, or dry powders. In this embodiment, the composition contains greater than 500 nM hypersensitive response elicitor polypeptide or protein.

- 27 -

Although not required, this composition may contain additional additives including fertilizer, insecticide, fungicide, nematocide, and mixtures thereof. Suitable fertilizers include $(\text{NH}_4)_2\text{NO}_3$. An example of a suitable insecticide is Malathion. Useful fungicides include Captan.

5 Other suitable additives include buffering agents, wetting agents, coating agents, and abrading agents. These materials can be used to facilitate the process of the present invention. In addition, the hypersensitive response elicitor polypeptide or protein can be applied to plant seeds with other conventional seed formulation and treatment materials, including clays and polysaccharides.

10 In the alternative embodiment of the present invention involving the use of transgenic plants and transgenic seeds, a hypersensitive response elicitor polypeptide or protein need not be applied topically to the plants or seeds. Instead, transgenic plants transformed with a DNA molecule encoding a hypersensitive response elicitor polypeptide or protein are produced according to procedures well
15 known in the art.

The vector described above can be microinjected directly into plant cells by use of micropipettes to transfer mechanically the recombinant DNA. Crossway, Mol. Gen. Genetics, 202:179-85 (1985), which is hereby incorporated by reference. The genetic material may also be transferred into the plant cell using
20 polyethylene glycol. Krens, et al., Nature, 296:72-74 (1982), which is hereby incorporated by reference.

Another approach to transforming plant cells with a gene which imparts resistance to pathogens is particle bombardment (also known as biolistic transformation) of the host cell. This can be accomplished in one of several ways.
25 The first involves propelling inert or biologically active particles at cells. This technique is disclosed in U.S. Patent Nos. 4,945,050, 5,036,006, and 5,100,792, all to Sanford et al., which are hereby incorporated by reference. Generally, this procedure involves propelling inert or biologically active particles at the cells under conditions effective to penetrate the outer surface of the cell and to be incorporated within the
30 interior thereof. When inert particles are utilized, the vector can be introduced into the cell by coating the particles with the vector containing the heterologous DNA. Alternatively, the target cell can be surrounded by the vector so that the vector is

- 28 -

carried into the cell by the wake of the particle. Biologically active particles (e.g., dried bacterial cells containing the vector and heterologous DNA) can also be propelled into plant cells.

Yet another method of introduction is fusion of protoplasts with other
5 entities, either minicells, cells, lysosomes or other fusible lipid-surfaced bodies. Fraley, et al., Proc. Natl. Acad. Sci. USA, 79:1859-63 (1982), which is hereby incorporated by reference.

The DNA molecule may also be introduced into the plant cells by electroporation. Fromm et al., Proc. Natl. Acad. Sci. USA, 82:5824 (1985), which is
10 hereby incorporated by reference. In this technique, plant protoplasts are electroporated in the presence of plasmids containing the expression cassette. Electrical impulses of high field strength reversibly permeabilize biomembranes allowing the introduction of the plasmids. Electroporated plant protoplasts reform the cell wall, divide, and regenerate.

Another method of introducing the DNA molecule into plant cells is to
15 infect a plant cell with *Agrobacterium tumefaciens* or *A. rhizogenes* previously transformed with the gene. Under appropriate conditions known in the art, the transformed plant cells are grown to form shoots or roots, and develop further into plants. Generally, this procedure involves inoculating the plant tissue with a
20 suspension of bacteria and incubating the tissue for 48 to 72 hours on regeneration medium without antibiotics at 25-28°C.

Agrobacterium is a representative genus of the gram-negative family Rhizobiaceae. Its species are responsible for crown gall (*A. tumefaciens*) and hairy
25 root disease (*A. rhizogenes*). The plant cells in crown gall tumors and hairy roots are induced to produce amino acid derivatives known as opines, which are catabolized only by the bacteria. The bacterial genes responsible for expression of opines are a convenient source of control elements for chimeric expression cassettes. In addition, assaying for the presence of opines can be used to identify transformed tissue.

Heterologous genetic sequences can be introduced into appropriate
30 plant cells, by means of the Ti plasmid of *A. tumefaciens* or the Ri plasmid of *A. rhizogenes*. The Ti or Ri plasmid is transmitted to plant cells on infection by

- 29 -

Agrobacterium and is stably integrated into the plant genome. J. Schell, Science, 237:1176-83 (1987), which is hereby incorporated by reference.

After transformation, the transformed plant cells must be regenerated.

Plant regeneration from cultured protoplasts is described in Evans et al., Handbook of Plant Cell Cultures, Vol. 1: (MacMillan Publishing Co., New York, 1983); and Vasil I.R. (ed.), Cell Culture and Somatic Cell Genetics of Plants, Acad. Press, Orlando, Vol. I, 1984, and Vol. III (1986), which are hereby incorporated by reference.

It is known that practically all plants can be regenerated from cultured cells or tissues, including but not limited to, all major species of sugarcane, sugar beets, cotton, fruit trees, and legumes.

Means for regeneration vary from species to species of plants, but generally a suspension of transformed protoplasts or a petri plate containing transformed explants is first provided. Callus tissue is formed and shoots may be induced from callus and subsequently rooted. Alternatively, embryo formation can be induced in the callus tissue. These embryos germinate as natural embryos to form plants. The culture media will generally contain various amino acids and hormones, such as auxin and cytokinins. It is also advantageous to add glutamic acid and proline to the medium, especially for such species as corn and alfalfa. Efficient regeneration will depend on the medium, on the genotype, and on the history of the culture. If these three variables are controlled, then regeneration is usually reproducible and repeatable.

After the expression cassette is stably incorporated in transgenic plants, it can be transferred to other plants by sexual crossing. Any of a number of standard breeding techniques can be used, depending upon the species to be crossed.

Once transgenic plants of this type are produced, the plants themselves can be cultivated in accordance with conventional procedure with the presence of the gene encoding the hypersensitive response elicitor resulting in disease resistance, enhanced plant growth, and/or control of insects on the plant. Alternatively, transgenic seeds are recovered from the transgenic plants. These seeds can then be planted in the soil and cultivated using conventional procedures to produce transgenic plants. The transgenic plants are propagated from the planted transgenic seeds under

- 30 -

conditions effective to impart disease resistance to plants, to enhance plant growth, and/or to control insects. While not wishing to be bound by theory, such disease resistance, growth enhancement, and/or insect control may be RNA mediated or may result from expression of the elicitor polypeptide or protein.

5 When transgenic plants and plant seeds are used in accordance with the present invention, they additionally can be treated with the same materials as are used to treat the plants and seeds to which a hypersensitive response elicitor polypeptide or protein is applied. These other materials, including hypersensitive response elicitors, can be applied to the transgenic plants and plant seeds by the above-noted procedures,
10 including high or low pressure spraying, injection, coating, and immersion. Similarly, after plants have been propagated from the transgenic plant seeds, the plants may be treated with one or more applications of the hypersensitive response elicitor to impart disease resistance, enhance growth, and/or control insects. Such plants may also be treated with conventional plant treatment agents (e.g., insecticides, fertilizers, etc.).

15 Another aspect of the present invention is to utilize the subject elicitor proteins or polypeptides to design molecules that will inactivate, destroy, or bind to these proteins or polypeptides. Since these elicitors are found in plant pathogens, particularly *Erwinia amylovora*, the pathogens themselves can be neutralized by the designed molecules so that disease and/or hypersensitive response is prevented or
20 altered. Examples of disease preventing molecules are antibodies, such as monoclonal or polyclonal antibodies, raised against the elicitor proteins or polypeptides of the present invention or binding portions thereof. Other examples of disease preventing molecules include antibody fragments, half-antibodies, hybrid derivatives, probes, and other molecular constructs.

25 Monoclonal antibody production may be effected by techniques which are well-known in the art. Basically, the process involves first obtaining immune cells (lymphocytes) from the spleen of a mammal (e.g., mouse) which has been previously immunized, either *in vivo* or *in vitro*, with the antigen of interest (e.g., an elicitor protein or polypeptide of the present invention or binding portions thereof).
30 The antibody-secreting lymphocytes are then fused with (mouse) myeloma cells or transformed cells, which are capable of replicating indefinitely in cell culture, thereby producing an immortal, immunoglobulin-secreting cell line. The resulting fused cells,

- 31 -

or hybridomas, are cultured, and the resulting colonies screened for the production of the desired monoclonal antibodies. Colonies producing such antibodies are cloned, and grown either *in vivo* or *in vitro* to produce large quantities of antibody. A description of the theoretical basis and practical methodology of fusing such cells is set forth in Kohler and Milstein, Nature 256:495 (1975), which is hereby incorporated by reference.

Mammalian lymphocytes are immunized by *in vivo* immunization of the animal (e.g., a mouse) with the elicitor proteins or polypeptides of the present invention or binding portions thereof. Such immunizations are repeated as necessary at intervals of up to several weeks to obtain a sufficient titer of antibodies. Following the last antigen boost, the animals are sacrificed and spleen cells removed.

Fusion with mammalian myeloma cells or other fusion partners capable of replicating indefinitely in cell culture is effected by standard and well-known techniques, for example, by using polyethylene glycol ("PEG") or other fusing agents (See Milstein and Kohler, Eur. J. Immunol. 6:511 (1976), which is hereby incorporated by reference). This immortal cell line, which is preferably murine, but may also be derived from cells of other mammalian species, including but not limited to rats, is selected to be deficient in enzymes necessary for the utilization of certain nutrients, to be capable of rapid growth, and to have good fusion capability. Many such cell lines are known to those skilled in the art, and others are regularly described.

Procedures for raising polyclonal antibodies are also well known. Typically, such antibodies can be raised by administering the elicitor proteins or polypeptides of the present invention or binding portions thereof subcutaneously to New Zealand white rabbits which have first been bled to obtain pre-immune serum. The antigens can be injected at a total volume of 100 μ l per site at six different sites. Each injected material will contain synthetic surfactant adjuvant pluronic polyols, or pulverized acrylamide gel containing the protein or polypeptide after SDS-polyacrylamide gel electrophoresis. The rabbits are then bled two weeks after the first injection and periodically boosted with the same antigen three times every six weeks. A sample of serum is then collected 10 days after each boost. Polyclonal antibodies are then recovered from the serum by affinity chromatography using the

- 32 -

corresponding antigen to capture the antibody. Ultimately, the rabbits are euthenized with pentobarbital 150 mg/Kg IV. This and other procedures for raising polyclonal antibodies are disclosed in E. Harlow, et. al., editors, Antibodies: A Laboratory Manual (1988), which is hereby incorporated by reference.

5 In addition to utilizing whole antibodies, the processes of the present invention encompass use of binding portions of such antibodies. Such binding portions include Fab fragments, F(ab')₂ fragments, and Fv fragments. These antibody fragments can be made by conventional procedures, such as proteolytic fragmentation procedures, as described in J. Goding, Monoclonal Antibodies: Principles and Practice, pp. 98-118 (N.Y. Academic Press 1983), which is hereby incorporated by
10 reference.

 Alternatively, the processes of the present invention can utilize probes or ligands found either in nature or prepared synthetically by recombinant DNA procedures or other biological or molecular procedures. Suitable probes or ligands
15 are molecules which bind to the elicitor proteins or polypeptides of the present invention or binding portions thereof.

 Avirulence (*avr*) genes (see Vivian, A., et al, Microbiology, 143:693-704 (1997); Leach, J. E., et al., Annu. Rev. Phytopathol., 34:153-179 (1996); Dangl, J. L. "Bacterial Pathogenesis of Plants and Animals: Molecular and Cellular
20 Mechanisms," in Current Topics in Microbiology and Immunology, Dangl, J. L., ed. (Springer, Berlin), Vol. 192, pp. 99-118 (1994), which are hereby incorporated by reference) generate signals that trigger defense responses leading to disease resistance in plants with corresponding resistance (*R*) genes. Typically, *avr* genes are isolated by expressing a cosmid library from one pathogen in another pathogen and screening
25 for narrowed host range. *avr* genes traditionally have been considered as negative determinants of host specificity at the race-cultivar level, but some, including the *avrE* locus from the bacterial speck pathogen *Pseudomonas syringae* pathovar (pv.) tomato (Kobayashi, D. Y., et al., Proc. Natl. Acad. Sci. USA, 86:157-61 (1989), which is hereby incorporated by reference), restrict host range at the pathovar-species or
30 species-species level (Whalen, M. C., et al., Proc. Natl. Acad. Sci. USA, 85:6743-47 (1988); Swarup, S., et al., Mol. Plant-Microbe Interact., 5:204-13 (1992), which are hereby incorporated by reference). Many *avr* genes, including *avrE*, are Hrp

- 33 -

regulated. *avrE* and *avrPphE* (Mansfield, J., et al., Mol. Plant-Microbe Interact., 7:726-39 (1994), which is hereby incorporated by reference) are physically linked to *hrp* genes.

When expressed *in trans*, the *avrE* locus renders *P. syringae* pv. glycinea, which causes bacterial blight of soybean, avirulent in all cultivars (Lorang, J. M., et al., Mol. Plant-Microbe Interact., 8:49-57 (1995), which is hereby incorporated by reference). The locus comprises two convergent transcription units, one preceded by a putative σ^{54} promoter and the other by a *hrp* box, a sequence found upstream of many *hrp* and *avr* genes that are positively regulated by the alternate
5 sigma factor HrpL (Innes, R. W., et al., J. Bacteriol., 175:4859-69 (1993); Shen, H., et al., J. Bacteriol., 175:5916-24 (1993); Xiao, Y., et al., J. Bacteriol., 176:3089-91 (1994), which are hereby incorporated by reference). Expression of both transcripts require HrpL. The *avrE* locus contributes quantitatively to the virulence in tomato leaves of *P. syringae* pv. tomato strain PT23, but not of strain DC3000 (Lorang, J. M.,
10 et al., Mol. Plant-Microbe Interact., 8:49-57 (1995); Lorang, J. M., et al., Mol. Plant-Microbe Interact. 7:508-515 (1994)).

Thus, *avr* genes in plant pathogens bind to disease resistance genes in plants which are not susceptible to that pathogen. In view of the homology of the DNA molecules of the present invention to *avr* genes in plant pathogens, these DNA
20 molecules can be used to identify corresponding plant disease resistance genes. Such identification is carried out by traditional plant breeding techniques in which a pathogen carrying the *avr* gene is inoculated to plants in screening to track inheritance or identify disruption of the resistance. Once identified, the resistance gene can be isolated by either of two approaches that have proved successful in recent years (see
25 Staskawicz et al., Science, 68:661-67 (1995)). These are positional or map-based cloning and insertional mutagenesis or transposon tagging. Because there may be no *DspE*-insensitive cultivars (susceptible to *Pseudomonas* harboring *dspE*; each of four soybean cultivars tested responded to *dspE*), map-based cloning (which requires crosses between susceptible and resistant lines to identify the position of the
30 resistance gene relative to other genes) may not be feasible. The preferred approach would more likely involve insertional mutagenesis, using the *dspE* gene or protein in

- 34 -

screens to identify lines which had lost the the product of *dspE* due to transposon tagging of the corresponding resistance gene.

EXAMPLES

5

Example 1 - Recombinant DNA techniques.

Isolation of DNA, restriction enzyme digests, ligation, transformation of *Escherichia coli*, and construction of and colony hybridization to screen a *P. syringae* pv. tomato DC3000 genomic library were performed as described by
10 Sambrook, et al. (Sambrook, J., et al., Molecular cloning: A Laboratory manual, (Cold Spring Harbor Laboratory, Cold Spring Harbor, NY) (1989), which is hereby incorporated by reference). The library was constructed using pCPP47 (Bauer, D. W., et al., Mol. Plant-Microbe Interact., 10:369-379 (1997), which is hereby incorporated by reference). Except where noted, *E. coli* DH5 and *E. coli* DH5 α were used as hosts
15 for DNA clones, and pBluescript or pBC plasmids (Stratagene, La Jolla, CA) were used as vectors. *E. amylovora* was transformed by electroporation as described (Bauer, D. W. in "Molecular Genetics of Pathogenicity of *Erwinia amylovora*: Techniques, Tools and Their Applications", (Ph. D. Thesis), Cornell University, Ithaca, NY (1990), which is hereby incorporated by reference). Plasmids were
20 mobilized into *E. amylovora* and *P. syringae* using pRK2013 (Figurski, D., et al., Proc. Natl. Acad. Sci. USA 76:1648-1652 (1979), which is hereby incorporated by reference).

Example 2 - Nucleotide sequencing and analysis.

25

The nucleotide sequence of the *dsp* region of *E. amylovora* strain Ea321 was determined using subclones of pCPP430 (Beer, S. V., et al., in Advances in Molecular Genetics of Plant-Microbe Interactions, Hennecke, H., et al.. eds. (Kluwer Academic Publishers, Dordrecht, The Netherlands), pp. 53-60 (1991), which is hereby incorporated by reference). The nucleotide sequence of the *avrE* locus was
30 determined using subclones of pCPP2357, a clone selected from a *P. syringae* pv. tomato DC3000 genomic cosmid library based on hybridization with the *hrpRS* operon of *P. syringae* pv. *syringae*, and the finding, based on partial sequencing, that it contained the *avrE* locus. Nucleotide sequencing was performed by the Cornell

- 35 -

Biotechnology Sequencing Facility on a Model 377 Sequencer (Perkin Elmer/Applied Biosystems Division, Foster City, CA). Sequence assembly, analysis, and comparisons were performed using the programs of the GCG software package, version 7.1 (Genetics Computer Groups, Inc., Madison, WI) and DNASTAR (DNASTAR, Inc., Madison, WI). Database searches were performed using BLAST (Altschul, S. F., et al., Proc. Nat. Acad. Sci. USA, 87:5509-5513 (1990) which is hereby incorporated by reference).

Example 3 - Expression of DspE and DspE' in E. coli.

10 The dspE operon was cloned in two pieces into pCPP50, a derivative of PINIII¹¹³-A2 (Duffaud, G. D., et al. in Methods in Enzymology, Wu, R., et al., eds. (Academic Press, New York), 153:492-50 (1987), which is hereby incorporated by reference) with an expanded polylinker, yielding pCPP1259. Expression in pCPP1259 is driven by the *Ipp* promoter of *E. coli*, under the control of the *lac* operator. An intermediate clone, pCPP1244, extending from the start of the operon to the *Bam*HI site in the middle of *dspE*, also was isolated. *E. coli* DH5 α strains containing pCPP1259 and pCPP1244 were grown in LB at 37°C to an OD₆₂₀ of 0.3. Isopropylthio- β -D-galactoside then was added to 1 mM, and the cells further incubated until reaching an OD₆₂₀ of 0.5. Cells were concentrated two-fold, lysed and subjected to SDS-PAGE as previously described (Sambrook, J., et al., Molecular cloning: A Laboratory Manual (Cold Spring Harbor Laboratory, Cold Spring Harbor, NY) (1989), which is hereby incorporated by reference), using 7.5% acrylamide. Cells containing pCPP50 were included for comparison. Proteins were visualized by Coomassie staining.

25 **Example 4 - Deletion mutagenesis of dspE.**

1554 bp were deleted from the 5' *Hind*III-*Bam*HI fragment of *dspE* in pCPP1237 using unique *Stu*I and *Sma*I sites. The mutagenized clone then was inserted into the suicide vector pKNG101 (Kaniga, K., et al., Gene, 109:137-42 (1991), which is hereby incorporated by reference) using *E. coli* SM10 λ pir as a host, yielding pCPP1241. The mutation, designated Δ 1554, then was transferred into *E. amylovora* strains using marker eviction as described previously (Bogdanove, A. J.,

- 36 -

et al., J. Bacteriol., 178:1720-30 (1996), which is hereby incorporated by reference). 1521 bp were deleted from the 3' *HindIII* fragment of *dspE* in pCPP1246 using two *BstEII* sites blunted with Klenow fragment. This mutation, $\Delta 1521$, was transferred into *E. amylovora* strains as above.

5

Example 5 - Pathogenicity assays.

For *E. amylovora* strains, cell suspensions of 5×10^8 colony-forming units (cfu) per ml were pipetted into wells cut in immature Bartlett pear fruit, or stabbed into Jonamac apple and cotoneaster shoots, and assays carried out as described previously (Beer, S. V., in Methods in Phytobacteriology, Klement, Z., et al., eds. (Adadémiai Kiadó, Budapest), pp. 373-374 (the "1990"); Aldwinckle, H. S., et al., Phytopathology, 66:1439-44 (1976), which are hereby incorporated by reference). For *P. syringae* pv. *glycinea*, panels of primary leaves of 2-week-old soybean seedlings (*Glycine max*, cultivar Norchief) were infiltrated with bacterial suspensions of 8×10^5 cfu/ml as for the HR assay, below. Plants were then covered with clear plastic bags and incubated under fluorescent lights (16 hr/day) at 22°C for 5-7 days. Leaves were scored for necrosis and chlorosis.

Example 6 - HR assays.

Tobacco leaf panels (*Nicotiana tabacum* L. 'xanthi') were infiltrated with bacterial cell suspensions as described previously (Wei, Z. M., et al., Science, 257:85-88 (1992); Bauer, D. W., et al., Mol. Plant-Microbe Interact., 4:493-99 (1991), which are hereby incorporated by reference). Primary leaves of 2-week-old soybean seedlings (secondary leaves emerging) were infiltrated with bacterial cell suspensions as for tobacco. Plants were scored for HR (tissue collapse) after 24-48 hr on the laboratory bench. *E. amylovora* strains were suspended in 5 mM KPO₄ buffer, pH 6.8, and *P. syringae* strains in 10 mM MgCl₂.

Example 7 - GUS assays.

Cells were 1.) grown in LB to an OD₆₂₀ of 0.9-1.0; 2.) grown in LB to an OD₆₂₀ of 0.5, then washed and resuspended in a *hrp*-gene-inducing minimal medium (Hrp MM; Huynh, T. V., et al., Science, 345:1374-77 (1989), which is

30

- 37 -

hereby incorporated by reference) to an OD₆₂₀ of 0.2 and incubated at 21° C for 36 hr to a final OD₆₂₀ of 0.9-1.0; or 3.) grown in LB to an OD₆₂₀ of 0.5, washed and concentrated 2-fold in 5 mM KPO₄ buffer, pH 6.8, and then transferred to freshly cut wells in pear halves and incubated as for the pathogenicity assay for 36 hr. Cells were

5 assayed for β-glucuronidase (GUS) activity essentially according to Jefferson (Jefferson, R. A., Plant Molecular Biology Reporter, 5:387-405 (1987), which is hereby incorporated by reference). For the cells in LB or Hrp MM, 50 μl were mixed with 200 μl GUS extraction buffer (50 mM NaHPO₄, pH 7.0, 10 mM β-mercaptoethanol, 10 mM Na₂EDTA, 0.1% sodium lauryl sarcosine, 0.1% Triton

10 X-100) containing 2 mM 4-methylumbelliferyl β-D-glucuronide as substrate and incubated at 37° C for 100 min. For cells in pear fruit, the tissue surrounding the well was excised using a #4 cork borer and homogenized in 5 mM KPO₄ buffer, pH 6.8. 200 μl of homogenate was mixed with 800 μl of GUS extraction buffer with substrate and incubated as above. Reactions were stopped by adding Na₂CO₃ to a final

15 concentration of 0.2 M in a total volume of 2 ml. Fluorescence was measured using a TKO 100 Mini-Fluorometer (Hoefer Scientific Instruments, San Francisco, CA). For all samples, cell concentration was estimated by dilution plating, and fluorometric readings were converted to pmole of substrate hydrolyzed/10⁸ cfu/min, after Miller (Miller, J. H., A Short Course in Bacterial Genetics: A Laboratory Manual and

20 Handbook for Escherichia coli and Related Bacteria (Cold Spring Harbor Laboratory Press, Plainview, NY) (1992), which is hereby incorporated by reference).

Example 8 - The “disease-specific” (dsp) region of *E. amylovora* consists of a 6.6 kb, two-gene operon.

25 Mapping of previous transposon insertions (Steinberger, E. M., et al., Mol. Plant-Microbe Interact., 1:135-44 (1988), which is hereby incorporated by reference) that abolish pathogenicity but not HR-eliciting ability confirmed the presence of the “disease specific” (dsp) region downstream of the *hrpN* gene in strain Ea321 as reported in strain CFBP1430 (Barny, A. M., et al., Mol. Microbiol., 4:777-

30 86 (1990), which is hereby incorporated by reference). The sequence of approximately 15 kb of DNA downstream of *hrpN* from Ea321 was determined, revealing several open reading frames (ORFs' Fig. 1). One ORF, in an apparent 6.6 kb operon with a

- 38 -

smaller ORF, spanned the area to which the *dsp* insertions mapped. These two ORFs were designated *dspE* and *dspF*, and the operon, *dspE*. *dspE* is preceded (beginning 70 bp upstream of the initiation codon) by the sequence GGAACCN₁₅CAACATAA, which matches the HrpL-dependent promoter consensus sequence or “*hrp*box” of *E. amylovora* (Kim, J. H., et al., J. Bacteriol., 179:1690-97 (1997); Kim, J. H., et al., J. Bacteriol., 179:1690-97 (1997), which are hereby incorporated by reference) and strongly resembles the *hrp* box of *P. syringae* *hrp* and *avr* genes (Xiao, Y., et al., J. Bacteriol., 176:3089-91 (1994), which is hereby incorporated by reference). Immediately downstream of *dspF* is A/T-rich DNA, followed by an ORF (ORF7) highly similar to the *Salmonella typhimurium* gene *spvR*, a member of the *lysR* family of regulatory genes (Caldwell, A. L. & Gulig, P. A., J. Bacteriol., 173:7176-85 (1991), which is hereby incorporated by reference). Immediately upstream of the *dspE* operon is a Hrp-regulated gene, *hrpW*, encoding a novel harpin.

The deduced product of *dspE* contains 1838 amino acid residues and is hydrophilic. The predicted molecular weight, 198 kD, was confirmed by expression in *E. coli* (Fig. 2). Expression of an intermediate clone containing only the 5' half of *dspE* yielded a protein of corresponding predicted mobility, suggesting that the N-terminal half of the protein might form an independently stable domain. *DspF*, predicted to be 16 kD, acidic (pI, 4.45), and predominantly α -helical, with amphipathic α helices in its C-terminus, is physically similar to virulence factor chaperones of animal-pathogenic bacteria (Wattiau, P., et al., Mol. Microbiol., 20:255-62 (1996), which is hereby incorporated by reference).

Example 9 - *dspE* is required for fire blight.

Two in-frame deletions within *dspE* (Fig. 1) were made in Ea321 and Ea273 (low- and high-virulence strains, respectively). The first (Δ 1554) corresponds to amino acid residues G₂₀₃ to G₇₂₀, and the second (Δ 1521) to amino acid residues T₁₀₆₄ to V₁₅₇₀. Each deletion abolished the ability of both strains to cause fire blight when inoculated to immature pear fruit (Fig. 3), apple shoots, or cotoneaster shoots. Δ 1554 was complemented by a clone carrying only the overlapping 5' half of *dspE*, further suggesting that the N-terminus of the protein forms a stable domain (Figs. 1 and 3).

Example 10 - The *dspE* operon contributes quantitatively and in a strain-dependent fashion to HR elicitation by *E. amylovora* in tobacco and is not required for HR elicitation by *E. amylovora* in soybean.

Transposon insertions in the *dsp* region reduce the ability of *E. amylovora* to elicit the HR in tobacco (Barney, A. M., et al., *Mol. Microbiol.*, 4:777-86 (1990), which is hereby incorporated by reference). Dilution series of suspensions of *dspE*Δ1554 mutant strains of Ea321 and Ea273 were infiltrated into tobacco leaves alongside their wild-type parents to assess the role of *dspE* in HR elicitation (Fig. 3). All strains were capable of eliciting the HR, but Ea321 *dspE*Δ1554, on a per-cell basis, was roughly one-tenth as effective as the wild-type in eliciting tissue collapse. There was no noticeable difference in HR-eliciting activity, however, between Ea273 and Ea273*dspE*Δ1554. Ea321*dspE*Δ1554 elicited wild-type HR in Acme, Centennial, Harasoy, and Norchief soybean leaves (Fig. 3).

Example 11 - The *dspE* operon is Hrp-regulated.

A promoterless *uidA* gene construct was cloned downstream of the *dspE* fragment in pCPP1241 that was used to introduce the Δ1554 mutation (Fig. 1) into wild-type strains of *E. amylovora* (this construct consists of a 3'-truncated *dspE* gene with the internal deletion). The resulting plasmid, pCPP1263, was mobilized into Ea321 and Ea273. Pathogenic strains, in which plasmid integration had preserved an intact copy of *dspE*, and non-pathogenic strains, in which the native copy of *dspE* had been mutated, were isolated. All strains were assayed for GUS activity in Luria Bertani medium (LB) and in Hrp MM, and pathogenic strains were assayed for activity in pear fruit. High levels of activity were obtained from strains incubated in Hrp MM and pear, but not LB. The level of expression in Hrp MM was equivalent to that of a *hrcV-uidA* fusion ("G73", Wei, et al., *J. Bacteriol.*, 177:6201-10 (1995), which is hereby incorporated by reference) used as a positive control. There were no significant differences in levels of expression of the *dspE-uidA* fusion in the wild-type and *dspE* mutant backgrounds, indicating that *dspE* likely is not autoregulated. Expression of the *dspE-uidA* fusion in *hrpL* mutants of Ea321 and

- 40 -

Ea273 in *hrp* MM was two orders of magnitude lower than that in HrpL + strains.

Data for Ea273 and derivatives are shown in Fig. 4.

Example 12 - *dspE* and *dspF* are homologous with genes in the *avrE* locus of *Pseudomonas syringae* pv. tomato.

A BLAST (Altschul, S. F., et al., J. Mol. Biol., 215:403-10 (1990), which is hereby incorporated by reference) search of the genetic databases revealed similarity of *dspE* to a partial sequence of the *avrE* locus of *P. syringae* pv. tomato (Lorang, J. M., et al., Mol. Plant-Microbe Interact., 8:49-57 (1995), which is hereby incorporated by reference). A cosmid library of *P. syringae* pv. tomato DC3000 genomic DNA was constructed, and a clone overlapping the *hrp* gene cluster and containing the *avrE* locus was isolated (pCPP2357). The complete nucleotide sequence of the *avrE* locus was determined, revealing the homolog of *dspE* (encoding a 195 kD, 1795 amino acid protein of 30% identity) alone in an operon previously designated transcription unit III, and a homolog of *dspF* (encoding a 14 kD, a 129 amino acid protein of 43% identity) at the end of the juxtaposed and opposing operon previously designated transcription unit IV (Fig. 1). These genes are designated *avrE* and *avrF*. The C-terminal half of the DspE and AvrE alignment (from V₈₄₅ of DspE) shows greater conservation (33% identity) than the N-terminal half (26% identity). AvrE contains a motif (aa residues A₄₅₀ to T₄₅₇) conserved in ATP- or GTP-binding proteins ("P-loop"; Saraste, M., et al., Trends Biochem. Sci., 15:430-34 (1990), which is hereby incorporated by reference). This motif is not conserved in DspE, however, and its functional significance in AvrE, if any, is unclear. Amino acid identities are distributed equally throughout the DspF and AvrF alignment, and AvrF shares the predicted physical characteristics of DspF. Upstream of *avrF*, competing the operon, is a 2.5 kb gene with no similarity to sequences in the genetic databases.

Example 13 - The *dspE* operon functions as an avirulence locus.

The *dspE* operon was cloned into pML 122 (Labes, M., et al., Gene, 89:37-46 (1990), which is hereby incorporated by reference) downstream of the *npIII* promoter, and this construct, pCPP1250, was mobilized into *P. syringae* pv. glycinea race 4. The resulting strain, but not a control strain containing pML 122, elicited the HR in soybean cultivars Acme, Centennial, Harasoy, and Norchief; in Norchief plants

- 41 -

incubated under conducive conditions, race 4 harboring pCPP1250 failed to cause symptoms of disease, while the control strain caused necrosis and chlorosis that spread from the point of inoculation (Fig. 5).

5 **Example 14 - *avrE* complements *dspE* mutations.**

Cosmid pCPP2357 was mobilized into Ea321 *dspE* mutant strains Δ 1554 and Δ 1521. The resulting transconjugants were pathogenic but low in virulence. Ea321*dspE* Δ 1521 carrying pCPP2357 with a transposon insertion in the *avrE* gene was non-pathogenic, demonstrating that the complementation observed
10 was *avrE*-specific (Figs. 1 and 5). The same results were observed for transconjugants of the Ea273*dspE* Δ 1521 mutant.

Over thirty bacterial *avr* genes have been discovered. The plethora of *avr* genes is thought to result from an "evolutionary tug-of-war" (Dangl, J. L., in Bacterial Pathogenesis of Plants and Animals: Molecular and Cellular Mechanisms
15 (Current Topics in Microbiology and Immunology), Dangl, J. L., ed. (Springer, Berlin), 192:99-118 (1994), which is hereby incorporated by reference), a reiterative process of selection, counterselection due to *R* genes, and modification or substitution of *avr* genes that was originally discerned by Flor, who hypothesized that "during their parallel evolution host and parasite developed complementary genic systems"
20 (Flor, H. H., Adv. Genet., 8:29-54 (1956), which is hereby incorporated by reference). However, only a few *avr* genes (including *avrE* in strain PT23) play detectable roles in virulence or pathogen fitness in their native genetic background (Lorang, J. M., et al., Mol. Plant-Microbe Interact., 7:508-15 (1994); Kearney, B., et al., Nature, 346:385-86 (1990); Swarup, S., et al., Phytopathology, 81:802-808 (1991); De Feyter, R. D., et al., Mol. Plant-Microbe Interact., 6:225-37 (1993); Ritter, C., et al., Mol.
25 Plant-Microbe Interact., 8:444-53 (1995), which are hereby incorporated by reference), and the selective force driving the maintenance in pathogen genomes of many of these host-range-limiting factors has remained a mystery. It is now clear, though, that several Avr proteins are delivered into plant cells by the Hrp pathway
30 (Gopalan, S., et al., Plant Cell, 8:1095-1105 (1996); Tang, X., et al., Science, 274:2060-63 (1996); Scofield, S. R., et al., Science, 274:2063-65 (1996); Leister, R. T., et al., Proc. Natl. Acad. Sci. USA, 93:15497-15502 (1996); Van Den Ackerveken,

- 42 -

G., et al., Cell, 87:1307-16 (1996), which are hereby incorporated by reference) and, therefore, are likely to be fundamentally virulence factors, which interact (directly, or indirectly through enzymatic products) with host targets to promote parasitism.

5 Mutation of such targets (selected because of reduced susceptibility) as well as the evolution of R proteins that recognize the Avr proteins would force the acquisition or evolution of new or modified Avr proteins and result in the proliferation of *avr* genes. Cumulatively, these co-evolutionary processes likely would drive a trend toward *avr* genes with quantitative and redundant effects in pathogenesis rather than critically important roles (Alfano, J. R., et al., Plant Cell, 8:1683-16988 (1996), which is hereby
10 incorporated by reference).

It has been found that the homologs *dspE* and *avrE* contribute to disease to dramatically different extents. The avirulence locus can substitute transgenerically for the pathogenicity operon, and that the avirulence function of *dspE* extends across pathogen genera as well. These findings support the hypothesis that
15 *avr* genes have a primary function in disease. Moreover, they support and expand the coevolutionary model for *avr* gene proliferation discussed above, and they have practical implications concerning the control of fire blight and other bacterial diseases of perennials.

One can predict from the model that the relative contribution to
20 pathogenicity of a particular factor would reflect, in part, the genetic history of the pathogen, specifically, the degree of co-evolution with its host(s). *dspE* is required for pathogenicity; *avrE* has a quantitative, strain-dependent, virulence phenotype. Consistent with the prediction, evolution of corresponding *R* genes and modification
of targets of pathogen virulence factors is likely to have occurred more often and to a
25 greater extent over time in the herbaceous hosts typically infected by *P. syringae* pathovars than in the woody hosts with which *E. amylovora* presumably evolved. Alternatively or additionally, acquisition of *dspE* (through evolution or horizontal transfer) by *E. amylovora* could have occurred relatively more recently than
acquisition of *avrE* by *P. syringae*, allowing less time for coevolution leading to
30 modification or the development of redundant function.

One could also hypothesize from the model that virulence factors may be conserved among pathogens, yet individually adapted to avoid detection on a

- 43 -

particular host. Preliminary results from Southern blot hybridizations suggest that *P. syringae* pv. *glycinea* harbors an *avrE* homolog, which, if functional, would support such a hypothesis. Similarly, homologs of the soybean cultivar-specific genes *avrA* and *avrD* from *P. syringae* pv. *tomato* exist in *P. syringae* pv. *glycinea* (Kobayashi, D. Y., et al., Proc. Natl. Acad. Sci. USA, 86:157-161 (1989), which is hereby incorporated by reference).

The homology and abilities of *dspE* and *avrE* to function transgenerically expand the model for *avr* gene proliferation. Major components of an evolution toward multifactor virulence could be procurement of genes encoding novel virulence factors from heterologous pathogens, and conservation of a functionally cosmopolitan virulence factor delivery system (and possibly conservation of a universal Hrp-pathway-targeting signal on the factors themselves) that would enable their deployment. Indeed, many *avr* genes are on plasmids and scattered in their distribution among pathogen strains (Dangl, J. L., in Bacterial Pathogenesis of Plants and Animals: Molecular and Cellular Mechanisms (Current Topics in Microbiology and Immunology), Dangl, J. L., ed. (Springer, Berlin), 192:99-118 (1994), which is hereby incorporated by reference), and individual *hrp* genes are conserved and even interchangeable (Arlat, M., et al., Mol. Plant-Microbe Interact., 4:593-601 (1991); Laby, R. J., et al., Mol. Plant-Microbe Interact., 5:412-19 (1992), which is hereby incorporated by reference). The presence of *dspE* and *avrE* in distinct genera suggests horizontal transfer of an ancestral locus, and, although *dspE* and *avrE* are homologous and *hrp*-linked, the transgeneric function of these genes suggests that the Hrp pathways in *E. amylovora* and *P. syringae* have remained insensitive to differences accrued in DspE and AvrE over evolution. It is predicted that even non-homologous Avr-like proteins will function across phytopathogenic bacterial genera.

It remains to be shown whether the avirulence function of the *dspE* locus is Hrp-pathway-dependent. This seems likely, and it will be important to determine the localization of the *dspE* and *dspF* gene products in the plant-bacterial interaction. The physical similarity of DspF (and AvrF) to chaperones required for type III secretion of virulence factors from animal-pathogenic bacteria (Wattiau, P., et al., Mol. Microbiol., 20:255-62 (1996), which is hereby incorporated by reference) is

- 44 -

intriguing and novel in phytopathogenic bacteria. The requirement of these chaperones appears to be due to a role other than targeting to the secretion pathway (Woestyn, S., et al., Mol. Microbiol., 20:1261-71 (1996), which is hereby incorporated by reference): chaperones may stabilize proteins, maintain proteins in an appropriate conformation for secretion, or prevent premature polymerization or association with other proteins. Perhaps, DspF binds to DspE (and AvrF to AvrE) and plays a similar role, which might be particularly important for the latter protein due to its large size and probable multidomain nature.

The *dspE* operon is the first described avirulence locus in *E. amylovora*. A homolog of *avrRxv* from *Xanthomonas campestris* (Whalen, M. C., et al., Proc. Natl. Acad. Sci. USA, 85:6743-47 (1988), which is hereby incorporated by reference) has been found near the *dspE* operon (Kim, J. F., in Molecular Characterization of a Novel Harpin and Two hrp Secretory Operons of Erwinia amylovora, and a hrp Operon of E. chrysanthemi (Ph.D. Thesis), Cornell University, Ithaca, NY (1997)). Monogenic (*R*-gene-mediated) resistance to fire blight has not been reported, but differential virulence of *E. amylovora* strains on apple cultivars has been observed (Norelli, J. L., et al., Phytopathology, 74:136-39 (1984), which is hereby incorporated by reference). Also, some strains of *E. amylovora* infect *Rubus* spp. and not pomaceous plants, and vice-versa (Starr, M. P., et al., Phytopathology, 41:915-19 (1951), which is hereby incorporated by reference). Whether the *dspE* operon and the *avrRxv* homolog or other potential elicitors play a role in these specificities should be determined.

Although the *dspE* operon triggers defense responses in soybean when expressed in *P. syringae* pv. *glycinea*, it is not required for the HR of soybean elicited by *E. amylovora*. Neither is *hrpN* required (Fig. 3). It is possible that *E. amylovora* must have one or the other, *dspE* or *hrpN*, to elicit the HR in soybean. It has been observed, however, that purified harpin does not elicit the HR in soybean, suggesting the alternative explanation that *E. amylovora* harbors another *avr* gene recognized by this plant.

Recognition of *E. amylovora* avirulence signals in soybean indicates the presence of one or more *R* genes that might be useful for engineering fire blight resistant apple and pear trees. *R*-gene-mediated resistance to the apple scab pathogen

- 45 -

Venturia inaequalis (Williams, E. B., et al., Ann. Rev. Phytopathol., 7:223-46 (1969), which is hereby incorporated by reference) and successful transformation of apple with attacin E for control of fire blight (Norelli, J. L., et al., Euphytica, 77:123-28 (1994), which is hereby incorporated by reference) attest the feasibility of such an approach. *R* gene-mediated resistance to apple scab has been overcome in the field (Parisi, L., et al., Phytopathology, 83:533-37 (1993), which is hereby incorporated by reference), but the requirement for *dspE* in disease favors relative durability of a *dspE*-specific *R* gene (Kearney, B. et al., Nature, 346:385-86 (1990), which is hereby incorporated by reference). Avirulence screening of *dspE* and other *E. amylovora* genes in pathogens of genetically tractable plants such as *Arabidopsis* could broaden the pool of candidate *R* genes and hasten their isolation. A similar approach could be used to isolate *R* genes effective against other diseases of woody plants. Furthermore, if the *dspE* operon is as widely conserved as is suggested by its homology with the *avrE* locus, a corresponding *R* gene could be effective against a variety of pathogens both of woody and herbaceous plants.

Native (non-denatured) DspE protein has not been produced in sufficient quantity to test its ability to elicit the HR (i.e. hypersensitive response) in a manner similar to hypersensitive response elicitors (i.e., by exogenous application). Therefore, no one has shown that *dspE* of *E. amylovora* elicits the HR when applied to plants as an isolated cell-free material. However, when the gene encoding the protein is transferred to another bacterium (along with the smaller *dspF* gene), e.g., *Pseudomonas syringae*, which ordinarily causes disease on certain plants, the recipient bacterium no longer causes disease but instead elicits the HR. The mechanism for this is not known for sure, but it is suspected to involve (and there is compelling evidence for) a mechanism in which the bacterial cell actually injects the DspE protein into the living plant cell, triggering the development of plant cell collapse (i.e. HR). Presumably, when the DspE protein is in the living plant cell, it might signal the plant to develop resistance to insects and pathogens.

Based on the similarity of the predicted physical characteristics of DspF to those of known chaperone proteins from animal pathogens, it is believed that this rather small protein is a chaperone of DspE. Chaperones in animal pathogens bind in the cytoplasm to specific proteins to be secreted. They seem to be required for

- 46 -

secretion of the proteins but are not themselves secreted. Evidence suggests that the chaperones are not involved directly in targeting the secreted proteins to the secretion apparatus. Instead, they may act to stabilize the proteins in the cytoplasm and/or prevent their premature aggregation or association with other proteins (e.g., bacterial proteins that direct transport through the host cell-membrane).

The *dspE* gene bears no similarity to known genes except *avrE*. Enzymatic function (i.e., one resulting in the production of a secondary molecule that elicits the HR) of *DspE* cannot be ruled out at present. In fact, one *avr* gene product is known to elicit HR indirectly by catalyzing synthesis of a diffusible elicitor molecule. However, the simplest explanation for the observed HR eliciting function of the *dspE* operon expressed in *Pseudomonas* species is that the protein encoded by the *dspE* gene is secreted from the bacterium and possibly transported into the plant cell, that there it triggers directly plant defense responses leading to the HR, and that this process is mediated by a specific resistance gene product that recognizes (acts as a receptor of) the *DspE* protein. Indeed, four *avr* genes that depend on the *Hrp* secretory apparatus to function when expressed in bacteria have been shown to cause HR when expressed transgenically within plant cells. One of these has been shown to encode a protein that directly interacts with the product of its corresponding resistance gene. Ultimately, whether *DspE* elicits plant defense responses from outside or inside the plant cell, directly or through a secondary molecule, must be determined in order to define practical applications of this protein and its encoding gene as a plant defense elicitor.

Although the invention has been described in detail for the purpose of illustration, it is understood that such detail is solely for that purpose, and variations can be made therein by those skilled in the art without departing from the spirit and scope of the invention which is defined by the following claims.

- 47 -

SEQUENCE LISTING

(1) GENERAL INFORMATION:

- (i) APPLICANT: Cornell Research Foundation, Inc.
- (ii) TITLE OF INVENTION: HYPERSENSITIVE RESPONSE ELICITOR FROM
ERWINIA AMYLOVORA, ITS USE, AND
ENCODING GENE
- (iii) NUMBER OF SEQUENCES: 5
- (iv) CORRESPONDENCE ADDRESS:
 - (A) ADDRESSEE: Nixon, Hargrave, Devans & Doyle LLP
 - (B) STREET: P.O. Box 1051, Clinton Square
 - (C) CITY: Rochester
 - (D) STATE: New York
 - (E) COUNTRY: U.S.A.
 - (F) ZIP: 14603
- (v) COMPUTER READABLE FORM:
 - (A) MEDIUM TYPE: Floppy disk
 - (B) COMPUTER: IBM PC compatible
 - (C) OPERATING SYSTEM: PC-DOS/MS-DOS
 - (D) SOFTWARE: PatentIn Release #1.0, Version #1.30
- (vi) CURRENT APPLICATION DATA:
 - (A) APPLICATION NUMBER:
 - (B) FILING DATE:
 - (C) CLASSIFICATION:
- (vii) PRIOR APPLICATION DATA:
 - (A) APPLICATION NUMBER: US 60/055,105
 - (B) FILING DATE: 06-AUG-1997
- (viii) ATTORNEY/AGENT INFORMATION:
 - (A) NAME: Goldman, Michael L.
 - (B) REGISTRATION NUMBER: 30,727
 - (C) REFERENCE/DOCKET NUMBER: 19603/1662
- (ix) TELECOMMUNICATION INFORMATION:
 - (A) TELEPHONE: (716) 263-1304
 - (B) TELEFAX: (716) 263-1600

(2) INFORMATION FOR SEQ ID NO:1:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 5517 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single

- 48 -

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

ATGGAATTAA AATCACTGGG AACTGAACAC AAGGCGGCAG TACACACAGC GGCGCACAAC	60
CCTGTGGGGC ATGGTGTTC CTTACAGCAG GGCAGCAGCA GCAGCAGCCC GCAAAATGCC	120
GCTGCATCAT TGGCGGCAGA AGGCAAAAAT CGTGGGAAAA TGCCGAGAAT TCACCAGCCA	180
TCTACTGCGG CTGATGGTAT CAGCGCTGCT CACCAGCAAA AGAAATCCTT CAGTCTCAGG	240
GGCTGTTTGG GGACGAAAAA ATTTTCCAGA TCGGCACCGC AGGGCCAGCC AGGTACCACC	300
CACAGCAAAG GGGCAACATT GCGCGATCTG CTGGCGCGGG ACGACGGCGA AACGCAGCAT	360
GAGGCGGCCG CGCCAGATGC GCGCGTTCG ACCCGTTCGG GCGCGTCAA ACGCCGCAAT	420
ATGGACGACA TGGCCGGGCG GCCAATGGTG AAAGGTGGCA GCGGCGAAGA TAAGGTACCA	480
ACGCAGCAAA AACGCATCA GCTGAACAAT TTTGGCCAGA TGCGCCAAAC GATGTTGAGC	540
AAAATGGCTC ACCCGGCTTC AGCCAACGCC GGCGATCGCC TGCAGCATTC ACCGCCGCAC	600
ATCCCGGGTA GCCACCACGA AATCAAGGAA GAACCGGTTG GCTCCACCAG CAAGGCAACA	660
ACGCCCCACG CAGACAGAGT GGAAATCGCT CAGGAAGATG ACGACAGCGA ATTCCAGCAA	720
CTGCATCAAC AGCGGCTGGC GCGCGAACGG GAAAATCCAC CGCAGCCGCC CAAACTCGGC	780
GTGCCACAC CGATTAGCGC CAGGTTTCAG CCCAACTGA CTGCGGTTGC GGAAAGCGTC	840
CTTGAGGGGA CAGATACCAC GCAGTCACCC CTTAAGCCGC AATCAATGCT GAAAGGAAGT	900
GGAGCCGGGG TAACGCCGCT GCGGTAACG CTGGATAAAG GCAAGTTGCA GCTGGCACCG	960
GATAATCCAC CCGCGCTCAA TACGTTGTTG AAGCAGACAT TGGGTAAAGA CACCCAGCAC	1020
TATCTGGCGC ACCATGCCAG CAGCGACGGT AGCCAGCATC TGCTGCTGGA CAACAAAGGC	1080
CACCTGTTTG ATATCAAAAG CACCGCCACC AGCTATAGCG TGCTGCACAA CAGCCACCCC	1140
GGTGAGATAA AGGGCAAGCT GGCGCAGGCG GGTACTGGCT CCGTCAGCGT AGACGGTAAA	1200
AGCGGCAAGA TCTCGCTGGG GAGCGGTACG CAAAGTCACA ACAAACAAT GCTAAGCCAA	1260
CCGGGGGAAG CGCACCGTTC CTATTAAACC GGCATTGTCG AGCATCCTGC TGGCGCAGCG	1320
CGGCGCAGG GCGAGTCAAT CCGCCTGCAT GACGACAAAA TTCATATCCT GCATCCGGAG	1380
CTGGGCGTAT GGCAATCTGC GGATAAAGAT ACCCACAGCC AGCTGTCTCG CCAGGCAGAC	1440
GGTAAGCTCT ATGCGCTGAA AGACAACCGT ACCCTGCAAA ACCTCTCCGA TAATAAATCC	1500
TCAGAAAAGC TGGTCGATAA AATCAAATCG TATTCCGTTG ATCAGCGGGG GCAGGTGGCG	1560
ATCCTGACGG ATACTCCCGG CCGCCATAAG ATGAGTATTA TGCCCTCGCT GGATGCTTCC	1620

- 49 -

CCGAGAGGCC ATATTTCCCT CAGCCTGCAT TTTGCCGATG CCCACCAGGG GTTATTGCAC	1680
GGGAAGTCGG AGCTTGAGGC ACAATCTGTC GCGATCAGCC ATGGGCGACT GGTGTGGCC	1740
GATAGCGAAG GCAAGCTGTT TAGCGCCGCC ATTCCGAAGC AAGGGGATGG AAACGAACTG	1800
AAAATGAAAG CCATGCCTCA GCATGCGCTC GATGAACATT TTGGTCATGA CCACCAGATT	1860
TCTGGATTTT TCCATGACGA CCACGGCCAG CTTAATGCGC TGGTGAAAAA TAACCTCAGG	1920
CAGCAGCATG CCTGCCCCGT GGGTAACGAT CATCAGTTTC ACCCGGCTG GAACCTGACT	1980
GATGCGCTGG TTATCGACAA TCAGCTGGGG CTGCATCATA CCAATCCTGA ACCGCATGAG	2040
ATTCTTGATA TGGGGCATT AGGCAGCCTG GCGTTACAGG AGGGCAAGCT TCACTATTTT	2100
GACCAGCTGA CCAAAGGGTG GACTGGCGCG GAGTCAGATT GTAAGCAGCT GAAAAAGGC	2160
CTGGATGGAG CAGCTTATCT ACTGAAAGAC GGTGAAGTGA AACGCCTGAA TATTAATCAG	2220
AGCACCTCCT CTATCAAGCA CGGAACGGAA AACGTTTTTT CGCTGCCGCA TGTGCGCAAT	2280
AAACCGGAGC CGGGAGATGC CCTGCAAGGG CTGAATAAAG ACGATAAGGC CCAGGCCATG	2340
GCGGTGATTG GGGTAAATAA ATACCTGGCG CTGACGGAAG AAGGGGACAT TCGCTCCTTC	2400
CAGATAAAAC CCGGCACCCA GCAGTTGGAG CGGCCGGCAC AAACCTCAG CCGCGAAGGT	2460
ATCAGCGGCG AACTGAAAGA CATTCATGTC GACCACAAGC AGAACCTGTA TGCCTTGACC	2520
CACGAGGGAG AGGTGTTTCA TCAGCCGCGT GAAGCCTGGC AGAATGGTGC CGAAAGCAGC	2580
AGCTGGCACA AACTGGCGTT GCCACAGAGT GAAAGTAAGC TAAAAAGTCT GGACATGAGC	2640
CATGAGCACA AACCGATTGC CACCTTTGAA GACGGTAGCC AGCATCAGCT GAAGGCTGGC	2700
GGCTGGCAGC CCTATGCGGC ACCTGAACGC GGGCCGCTGG CGGTGGGTAC CAGCGTTCA	2760
CAAACCGTCT TTAACCGACT AATGCAGGGG GTGAAAGGCA AGGTGATCCC AGGCAGCGGG	2820
TTGACGGTTA AGCTCTCGGC TCAGACGGGG GGAATGACCG GCGCCGAAGG GCGCAAGGTC	2880
AGCAGTAAAT TTTCCGAAAG GATCCGCGCC TATGCGTTCA ACCCAACAAT GTCCACGCCG	2940
CGACCGATTA AAAATGCTGC TTATGCCACA CAGCACGGCT GGCAGGGGCG TGAGGGGTG	3000
AAGCCGTGTG ACGAGATGCA GGGAGCGCTG ATTAAACAAC TGGATGCGCA TAACGTTCTG	3060
CATAACGCGC CACAGCCAGA TTTGCAGAGC AAACCTGAAA CTCTGGATTT AGGCGAACAT	3120
GGCGCAGAAT TGCTTAACGA CATGAAGCGC TTCCGCGACG AACTGGAGCA GAGTGCAACC	3180
CGTTCGGTGA CCGTTTTAGG TCAACATCAG GGAGTGCTAA AAAGCAACGG TGAAATCAAT	3240
AGCGAATTTA AGCCATCGCC CGGCAAGGCG TTGGTCCAGA GCTTTAACGT CAATCGCTCT	3300
GGTCAGGATC TAAGCAAGTC ACTGCAACAG GCAGTACATG CCACGCCGCC ATCCGCAGAG	3360
AGTAACTGC AATCCATGCT GGGGCACCTT GTCAGTGCCG GGGTGGATAT GAGTCATCAG	3420
AAGGGCGAGA TCCCGCTGGG CCGCCAGCGC GATCCGAATG ATAAAACGC ACTGACCAAA	3480

- 50 -

TCGCGTTTAA TTTTAGATAC CGTGACCATC GGTGAACTGC ATGAACTGGC CGATAAGGCG	3540
AAACTGGTAT CTGACCATAA ACCCGATGCC GATCAGATAA AACAGCTGCG CCAGCAGTTC	3600
GATACGCTGC GTGAAAAGCG GTATGAGAGC AATCCGGTGA AGCATTACAC CGATATGGGC	3660
TTCACCCATA ATAAGGCGCT GGAAGCAAAC TATGATGCGG TCAAAGCCTT TATCAATGCC	3720
TTTAAGAAAG AGCACCACGG CGTCAATCTG ACCACGCGTA CCGTACTGGA ATCACAGGGC	3780
AGTCCGGAGC TGGCGAAGAA GCTCAAGAAT ACGCTGTTGT CCCTGGACAG TGGTGAAAGT	3840
ATGAGCTTCA GCCGGTCATA TGGCGGGGGC GTCAGCACTG TCTTTGTGCC TACCCTTAGC	3900
AAGAAGGTGC CAGTTCGGT GATCCCCGGA GCCGGCATCA CGCTGGATCG CGCCTATAAC	3960
CTGAGCTTCA GTCGTACCAG CGGCGGATTG AACGTCAGTT TTGGCCGCGA CGGCGGGGTG	4020
AGTGGTAACA TCATGGTCGC TACCGGCCAT GATGTGATGC CCTATATGAC CGGTAAGAAA	4080
ACCACTGCAG GTAACGCCAG TGA CTGTTG AGCGCAAAC ATAAATCAG CCCGACTTG	4140
CGTATCGGCG CTGCTGTGAG TGGCACCTG CAAGGAACGC TACAAAACAG CCTGAAGTTT	4200
AAGCTGACAG AGGATGAGCT GCCTGGCTTT ATCCATGGCT TGACGCATGG CACGTTGACC	4260
CCGGCAGAAC TGTGCAAAA GGGGATCGAA CATCAGATGA AGCAGGGCAG CAACTGACG	4320
TTTAGCGTCG ATACCTCGGC AAATCTGGAT CTGCGTGCCG GTATCAATCT GAACGAAGAC	4380
GGCAGTAAAC CAAATGGTGT CACTGCCCGT GTTCTGCGG GGCTAAGTGC ATCGGCAAAC	4440
CTGGCCGCCG GCTCGCGTGA ACGCAGCACC ACCTCTGGCC AGTTTGGCAG CACGACTTCG	4500
GCCAGCAATA ACCGCCCAAC CTTCTCAAC GGGGTGCGCG CGGGTGCTAA CCTGACGGCT	4560
GCTTTAGGGG TTGCCCATC ATCTACGCAT GAAGGGAAC CGGTCGGGAT CTTCCCGGCA	4620
TTTACCTCGA CCAATGTTTC GGCAGCGCTG GCGCTGGATA ACCGTACCTC ACAGAGTATC	4680
AGCCTGGAAT TGAAGCGCGC GGAGCCGGTG ACCAGCAACG ATATCAGCGA GTTGACCTCC	4740
ACGCTGGGAA AACACTTTAA GGATAGCGCC ACAACGAAGA TGCTTGCCGC TCTCAAAGAG	4800
TTAGATGACG CTAAGCCCGC TGAACAAC TG CATATTTTAC AGCAGCATT CAGTGCAAAA	4860
GATGTCGTCG GTGATGAACG CTACGAGGCG GTGCGCAACC TGAAAAAACT GGTGATACGT	4920
CAACAGGCTG CGGACAGCCA CAGCATGGAA TTAGGATCTG CCAGTCACAG CACGACCTAC	4980
AATAATCTGT CGAGAATAAA TAATGACGGC ATTGTCGAGC TGCTACACAA ACATTTTCGAT	5040
GCGGCATTAC CAGCAAGCAG TGCCAAACGT CTTGGTGAAA TGATGAATAA CGATCCGGCA	5100
CTGAAAGATA TTATTAAGCA GCTGCAAAGT ACGCCGTTCA GCAGCGCCAG CGTGTCGATG	5160
GAGCTGAAAG ATGGTCTGCG TGAGCAGACG GAAAAAGCAA TACTGGACGG TAAGGTCGGT	5220
CGTGAAGAAG TGGGAGTACT TTTCCAGGAT CGTAACAAC TGCGTGTTAA ATCGGTCAGC	5280
GTCAGTCAGT CCGTCAGCAA AAGCGAAGGC TTCAATACCC CAGCGCTGTT ACTGGGGACG	5340

- 51 -

AGCAACAGCG CTGCTATGAG CATGGAGCGC AACATCGGAA CCATTAATTT TAAATACGGC 5400
 CAGGATCAGA ACACCCACG GCGATTTACC CTGGAGGGTG GAATAGCTCA GGCTAATCCG 5460
 CAGGTCGCAT CTGCGCTTAC TGATTTGAAG AAGGAAGGGC TGGAAATGAA GAGCTAA 5517

(2) INFORMATION FOR SEQ ID NO:2:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1838 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS:
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

Met	Glu	Leu	Lys	Ser	Leu	Gly	Thr	Glu	His	Lys	Ala	Ala	Val	His	Thr	1	5	10	15
Ala	Ala	His	Asn	Pro	Val	Gly	His	Gly	Val	Ala	Leu	Gln	Gln	Gly	Ser	20	25	30	
Ser	Ser	Ser	Ser	Pro	Gln	Asn	Ala	Ala	Ala	Ser	Leu	Ala	Ala	Glu	Gly	35	40	45	
Lys	Asn	Arg	Gly	Lys	Met	Pro	Arg	Ile	His	Gln	Pro	Ser	Thr	Ala	Ala	50	55	60	
Asp	Gly	Ile	Ser	Ala	Ala	His	Gln	Gln	Lys	Lys	Ser	Phe	Ser	Leu	Arg	65	70	75	80
Gly	Cys	Leu	Gly	Thr	Lys	Lys	Phe	Ser	Arg	Ser	Ala	Pro	Gln	Gly	Gln	85	90	95	
Pro	Gly	Thr	Thr	His	Ser	Lys	Gly	Ala	Thr	Leu	Arg	Asp	Leu	Leu	Ala	100	105	110	
Arg	Asp	Asp	Gly	Glu	Thr	Gln	His	Glu	Ala	Ala	Ala	Pro	Asp	Ala	Ala	115	120	125	
Arg	Leu	Thr	Arg	Ser	Gly	Gly	Val	Lys	Arg	Arg	Asn	Met	Asp	Asp	Met	130	135	140	
Ala	Gly	Arg	Pro	Met	Val	Lys	Gly	Gly	Ser	Gly	Glu	Asp	Lys	Val	Pro	145	150	155	160
Thr	Gln	Gln	Lys	Arg	His	Gln	Leu	Asn	Asn	Phe	Gly	Gln	Met	Arg	Gln	165	170	175	
Thr	Met	Leu	Ser	Lys	Met	Ala	His	Pro	Ala	Ser	Ala	Asn	Ala	Gly	Asp	180	185	190	
Arg	Leu	Gln	His	Ser	Pro	Pro	His	Ile	Pro	Gly	Ser	His	His	Glu	Ile	195	200	205	

- 52 -

Lys Glu Glu Pro Val Gly Ser Thr Ser Lys Ala Thr Thr Ala His Ala
 210 215 220
 Asp Arg Val Glu Ile Ala Gln Glu Asp Asp Asp Ser Glu Phe Gln Gln
 225 230 235 240
 Leu His Gln Gln Arg Leu Ala Arg Glu Arg Glu Asn Pro Pro Gln Pro
 245 250 255
 Pro Lys Leu Gly Val Ala Thr Pro Ile Ser Ala Arg Phe Gln Pro Lys
 260 265 270
 Leu Thr Ala Val Ala Glu Ser Val Leu Glu Gly Thr Asp Thr Thr Gln
 275 280 285
 Ser Pro Leu Lys Pro Gln Ser Met Leu Lys Gly Ser Gly Ala Gly Val
 290 295 300
 Thr Pro Leu Ala Val Thr Leu Asp Lys Gly Lys Leu Gln Leu Ala Pro
 305 310 315 320
 Asp Asn Pro Pro Ala Leu Asn Thr Leu Leu Lys Gln Thr Leu Gly Lys
 325 330 335
 Asp Thr Gln His Tyr Leu Ala His His Ala Ser Ser Asp Gly Ser Gln
 340 345 350
 His Leu Leu Leu Asp Asn Lys Gly His Leu Phe Asp Ile Lys Ser Thr
 355 360 365
 Ala Thr Ser Tyr Ser Val Leu His Asn Ser His Pro Gly Glu Ile Lys
 370 375 380
 Gly Lys Leu Ala Gln Ala Gly Thr Gly Ser Val Ser Val Asp Gly Lys
 385 390 395 400
 Ser Gly Lys Ile Ser Leu Gly Ser Gly Thr Gln Ser His Asn Lys Thr
 405 410 415
 Met Leu Ser Gln Pro Gly Glu Ala His Arg Ser Leu Leu Thr Gly Ile
 420 425 430
 Trp Gln His Pro Ala Gly Ala Ala Arg Pro Gln Gly Glu Ser Ile Arg
 435 440 445
 Leu His Asp Asp Lys Ile His Ile Leu His Pro Glu Leu Gly Val Trp
 450 455 460
 Gln Ser Ala Asp Lys Asp Thr His Ser Gln Leu Ser Arg Gln Ala Asp
 465 470 475 480
 Gly Lys Leu Tyr Ala Leu Lys Asp Asn Arg Thr Leu Gln Asn Leu Ser
 485 490 495
 Asp Asn Lys Ser Ser Glu Lys Leu Val Asp Lys Ile Lys Ser Tyr Ser
 500 505 510
 Val Asp Gln Arg Gly Gln Val Ala Ile Leu Thr Asp Thr Pro Gly Arg
 515 520 525
 His Lys Met Ser Ile Met Pro Ser Leu Asp Ala Ser Pro Glu Ser His
 530 535 540

- 53 -

Ile Ser Leu Ser Leu His Phe Ala Asp Ala His Gln Gly Leu Leu His
 545 550 555 560
 Gly Lys Ser Glu Leu Glu Ala Gln Ser Val Ala Ile Ser His Gly Arg
 565 570 575
 Leu Val Val Ala Asp Ser Glu Gly Lys Leu Phe Ser Ala Ala Ile Pro
 580 585 590
 Lys Gln Gly Asp Gly Asn Glu Leu Lys Met Lys Ala Met Pro Gln His
 595 600 605
 Ala Leu Asp Glu His Phe Gly His Asp His Gln Ile Ser Gly Phe Phe
 610 615 620
 His Asp Asp His Gly Gln Leu Asn Ala Leu Val Lys Asn Asn Phe Arg
 625 630 635 640
 Gln Gln His Ala Cys Pro Leu Gly Asn Asp His Gln Phe His Pro Gly
 645 650 655
 Trp Asn Leu Thr Asp Ala Leu Val Ile Asp Asn Gln Leu Gly Leu His
 660 665 670
 His Thr Asn Pro Glu Pro His Glu Ile Leu Asp Met Gly His Leu Gly
 675 680 685
 Ser Leu Ala Leu Gln Glu Gly Lys Leu His Tyr Phe Asp Gln Leu Thr
 690 695 700
 Lys Gly Trp Thr Gly Ala Glu Ser Asp Cys Lys Gln Leu Lys Lys Gly
 705 710 715 720
 Leu Asp Gly Ala Ala Tyr Leu Leu Lys Asp Gly Glu Val Lys Arg Leu
 725 730 735
 Asn Ile Asn Gln Ser Thr Ser Ser Ile Lys His Gly Thr Glu Asn Val
 740 745 750
 Phe Ser Leu Pro His Val Arg Asn Lys Pro Glu Pro Gly Asp Ala Leu
 755 760 765
 Gln Gly Leu Asn Lys Asp Asp Lys Ala Gln Ala Met Ala Val Ile Gly
 770 775 780
 Val Asn Lys Tyr Leu Ala Leu Thr Glu Lys Gly Asp Ile Arg Ser Phe
 785 790 795 800
 Gln Ile Lys Pro Gly Thr Gln Gln Leu Glu Arg Pro Ala Gln Thr Leu
 805 810 815
 Ser Arg Glu Gly Ile Ser Gly Glu Leu Lys Asp Ile His Val Asp His
 820 825 830
 Lys Gln Asn Leu Tyr Ala Leu Thr His Glu Gly Glu Val Phe His Gln
 835 840 845
 Pro Arg Glu Ala Trp Gln Asn Gly Ala Glu Ser Ser Ser Trp His Lys
 850 855 860

- 54 -

Leu Ala Leu Pro Gln Ser Glu Ser Lys Leu Lys Ser Leu Asp Met Ser
 865 870 875 880
 His Glu His Lys Pro Ile Ala Thr Phe Glu Asp Gly Ser Gln His Gln
 885 890 895
 Leu Lys Ala Gly Gly Trp His Ala Tyr Ala Ala Pro Glu Arg Gly Pro
 900 905 910
 Leu Ala Val Gly Thr Ser Gly Ser Gln Thr Val Phe Asn Arg Leu Met
 915 920 925
 Gln Gly Val Lys Gly Lys Val Ile Pro Gly Ser Gly Leu Thr Val Lys
 930 935 940
 Leu Ser Ala Gln Thr Gly Gly Met Thr Gly Ala Glu Gly Arg Lys Val
 945 950 955 960
 Ser Ser Lys Phe Ser Glu Arg Ile Arg Ala Tyr Ala Phe Asn Pro Thr
 965 970 975
 Met Ser Thr Pro Arg Pro Ile Lys Asn Ala Ala Tyr Ala Thr Gln His
 980 985 990
 Gly Trp Gln Gly Arg Glu Gly Leu Lys Pro Leu Tyr Glu Met Gln Gly
 995 1000 1005
 Ala Leu Ile Lys Gln Leu Asp Ala His Asn Val Arg His Asn Ala Pro
 1010 1015 1020
 Gln Pro Asp Leu Gln Ser Lys Leu Glu Thr Leu Asp Leu Gly Glu His
 1025 1030 1035 1040
 Gly Ala Glu Leu Leu Asn Asp Met Lys Arg Phe Arg Asp Glu Leu Glu
 1045 1050 1055
 Gln Ser Ala Thr Arg Ser Val Thr Val Leu Gly Gln His Gln Gly Val
 1060 1065 1070
 Leu Lys Ser Asn Gly Glu Ile Asn Ser Glu Phe Lys Pro Ser Pro Gly
 1075 1080 1085
 Lys Ala Leu Val Gln Ser Phe Asn Val Asn Arg Ser Gly Gln Asp Leu
 1090 1095 1100
 Ser Lys Ser Leu Gln Gln Ala Val His Ala Thr Pro Pro Ser Ala Glu
 1105 1110 1115 1120
 Ser Lys Leu Gln Ser Met Leu Gly His Phe Val Ser Ala Gly Val Asp
 1125 1130 1135
 Met Ser His Gln Lys Gly Glu Ile Pro Leu Gly Arg Gln Arg Asp Pro
 1140 1145 1150
 Asn Asp Lys Thr Ala Leu Thr Lys Ser Arg Leu Ile Leu Asp Thr Val
 1155 1160 1165
 Thr Ile Gly Glu Leu His Glu Leu Ala Asp Lys Ala Lys Leu Val Ser
 1170 1175 1180
 Asp His Lys Pro Asp Ala Asp Gln Ile Lys Gln Leu Arg Gln Gln Phe
 1185 1190 1195 1200

- 55 -

Asp Thr Leu Arg Glu Lys Arg Tyr Glu Ser Asn Pro Val Lys His Tyr
 1205 1210 1215
 Thr Asp Met Gly Phe Thr His Asn Lys Ala Leu Glu Ala Asn Tyr Asp
 1220 1225 1230
 Ala Val Lys Ala Phe Ile Asn Ala Phe Lys Lys Glu His His Gly Val
 1235 1240 1245
 Asn Leu Thr Thr Arg Thr Val Leu Glu Ser Gln Gly Ser Ala Glu Leu
 1250 1255 1260
 Ala Lys Lys Leu Lys Asn Thr Leu Leu Ser Leu Asp Ser Gly Glu Ser
 1265 1270 1275 1280
 Met Ser Phe Ser Arg Ser Tyr Gly Gly Gly Val Ser Thr Val Phe Val
 1285 1290 1295
 Pro Thr Leu Ser Lys Lys Val Pro Val Pro Val Ile Pro Gly Ala Gly
 1300 1305 1310
 Ile Thr Leu Asp Arg Ala Tyr Asn Leu Ser Phe Ser Arg Thr Ser Gly
 1315 1320 1325
 Gly Leu Asn Val Ser Phe Gly Arg Asp Gly Gly Val Ser Gly Asn Ile
 1330 1335 1340
 Met Val Ala Thr Gly His Asp Val Met Pro Tyr Met Thr Gly Lys Lys
 1345 1350 1355 1360
 Thr Ser Ala Gly Asn Ala Ser Asp Trp Leu Ser Ala Lys His Lys Ile
 1365 1370 1375
 Ser Pro Asp Leu Arg Ile Gly Ala Ala Val Ser Gly Thr Leu Gln Gly
 1380 1385 1390
 Thr Leu Gln Asn Ser Leu Lys Phe Lys Leu Thr Glu Asp Glu Leu Pro
 1395 1400 1405
 Gly Phe Ile His Gly Leu Thr His Gly Thr Leu Thr Pro Ala Glu Leu
 1410 1415 1420
 Leu Gln Lys Gly Ile Glu His Gln Met Lys Gln Gly Ser Lys Leu Thr
 1425 1430 1435 1440
 Phe Ser Val Asp Thr Ser Ala Asn Leu Asp Leu Arg Ala Gly Ile Asn
 1445 1450 1455
 Leu Asn Glu Asp Gly Ser Lys Pro Asn Gly Val Thr Ala Arg Val Ser
 1460 1465 1470
 Ala Gly Leu Ser Ala Ser Ala Asn Leu Ala Ala Gly Ser Arg Glu Arg
 1475 1480 1485
 Ser Thr Thr Ser Gly Gln Phe Gly Ser Thr Thr Ser Ala Ser Asn Asn
 1490 1495 1500
 Arg Pro Thr Phe Leu Asn Gly Val Gly Ala Gly Ala Asn Leu Thr Ala
 1505 1510 1515 1520

- 56 -

Ala Leu Gly Val	Ala His Ser Ser Thr	His Glu Gly Lys	Pro Val Gly
1525		1530	1535
Ile Phe Pro Ala Phe Thr	Ser Thr Asn Val Ser Ala Ala	Leu Ala Leu	
1540	1545	1550	
Asp Asn Arg Thr Ser Gln Ser	Ile Ser Leu Glu Leu Lys Arg	Ala Glu	
1555	1560	1565	
Pro Val Thr Ser Asn Asp	Ile Ser Glu Leu Thr Ser Thr	Leu Gly Lys	
1570	1575	1580	
His Phe Lys Asp Ser Ala Thr Thr	Lys Met Leu Ala Ala Leu Lys Glu		
1585	1590	1595	1600
Leu Asp Asp Ala Lys Pro Ala Glu Gln	Leu His Ile Leu Gln Gln His		
1605	1610	1615	
Phe Ser Ala Lys Asp Val Val Gly Asp	Glu Arg Tyr Glu Ala Val Arg		
1620	1625	1630	
Asn Leu Lys Lys Leu Val Ile Arg Gln Gln	Ala Ala Asp Ser His Ser		
1635	1640	1645	
Met Glu Leu Gly Ser Ala Ser His Ser Thr Thr	Tyr Asn Asn Leu Ser		
1650	1655	1660	
Arg Ile Asn Asn Asp Gly Ile Val Glu Leu Leu	His Lys His Phe Asp		
1665	1670	1675	1680
Ala Ala Leu Pro Ala Ser Ser Ala Lys Arg	Leu Gly Glu Met Met Asn		
1685	1690	1695	
Asn Asp Pro Ala Leu Lys Asp Ile Ile Lys Gln	Leu Gln Ser Thr Pro		
1700	1705	1710	
Phe Ser Ser Ala Ser Val Ser Met Glu Leu Lys	Asp Gly Leu Arg Glu		
1715	1720	1725	
Gln Thr Glu Lys Ala Ile Leu Asp Gly Lys Val	Gly Arg Glu Glu Val		
1730	1735	1740	
Gly Val Leu Phe Gln Asp Arg Asn Asn Leu Arg	Val Lys Ser Val Ser		
1745	1750	1755	1760
Val Ser Gln Ser Val Ser Lys Ser Glu Gly Phe	Asn Thr Pro Ala Leu		
1765	1770	1775	
Leu Leu Gly Thr Ser Asn Ser Ala Ala Met Ser	Met Glu Arg Asn Ile		
1780	1785	1790	
Gly Thr Ile Asn Phe Lys Tyr Gly Gln Asp Gln	Asn Thr Pro Arg Arg		
1795	1800	1805	
Phe Thr Leu Glu Gly Gly Ile Ala Gln Ala Asn	Pro Gln Val Ala Ser		
1810	1815	1820	
Ala Leu Thr Asp Leu Lys Lys Glu Gly Leu Glu	Met Lys Ser		
1825	1830	1835	

- 57 -

(2) INFORMATION FOR SEQ ID NO:3:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 420 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

```

ATGACATCGT CACAGCAGCG GGTGAAAGG TTTTACAGT ATTTCTCCGC CGGGTGTA   60
ACGCCCATAC ATCTGAAAGA CGGGGTGTGC GCCCTGTATA ACGAACAAGA TGAGGAGGCG   120
GCGGTGCTGG AAGTACCGCA ACACAGCGAC AGCCTGTTAC TACACTGCCG AATCATTGAG   180
GCTGACCCAC AAACITCAAT AACCTGTAT TCGATGCTAT TACAGCTGAA TTTGAAATG   240
GCGGCCATGC GCGGCTGTTG GCTGGCGCTG GATGAACTGC ACAACGTGCG TTTATGTTTT   300
CAGCAGTCGC TGGAGCATCT GGATGAAGCA AGTTTTAGCG ATATCGTTAG CGGCTTCATC   360
GAACATGCGG CAGAAGTGCG TGAGTATATA GCGCAATTAG ACGAGAGTAG CGCGGCATAA   420

```

(2) INFORMATION FOR SEQ ID NO:4:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 139 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS:
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

```

Met Thr Ser Ser Gln Gln Arg Val Glu Arg Phe Leu Gln Tyr Phe Ser
1           5           10          15

Ala Gly Cys Lys Thr Pro Ile His Leu Lys Asp Gly Val Cys Ala Leu
20          25          30

Tyr Asn Glu Gln Asp Glu Glu Ala Ala Val Leu Glu Val Pro Gln His
35          40          45

Ser Asp Ser Leu Leu Leu His Cys Arg Ile Ile Glu Ala Asp Pro Gln
50          55          60

Thr Ser Ile Thr Leu Tyr Ser Met Leu Leu Gln Leu Asn Phe Glu Met
65          70          75          80

Ala Ala Met Arg Gly Cys Trp Leu Ala Leu Asp Glu Leu His Asn Val
85          90          95

```

- 58 -

Arg Leu Cys Phe Gln Gln Ser Leu Glu His Leu Asp Glu Ala Ser Phe
100 105 110

Ser Asp Ile Val Ser Gly Phe Ile Glu His Ala Ala Glu Val Arg Glu
115 120 125

Tyr Ile Ala Gln Leu Asp Glu Ser Ser Ala Ala
130 135

(2) INFORMATION FOR SEQ ID NO:5:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 29 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

GGAACCNNNN NNNNNNNNNN NCAACATAA

29

WHAT IS CLAIMED:

1. An isolated DNA molecule encoding a hypersensitive response eliciting protein or polypeptide, wherein the isolated DNA molecule is selected from the group consisting of (a) a DNA molecule comprising a nucleotide sequence of SEQ. ID. Nos. 1 or 3, (b) a DNA molecule encoding a protein comprising an amino acid of SEQ. ID. Nos. 2 or 4, (c) a DNA molecule which hybridizes to a DNA molecule comprising a nucleotide sequence of SEQ. ID. Nos. 1 or 3, under stringent conditions, and (d) a DNA molecule complementary to DNA molecules (a), (b), and (c).
2. An isolated DNA molecule according to claim 1, wherein said DNA molecule is a DNA molecule comprising a nucleotide sequence of SEQ. ID. Nos. 1 or 3.
3. An isolated DNA molecule according to claim 1, wherein said DNA molecule is a DNA molecule encoding protein comprising an amino acid of SEQ. ID. Nos. 2 or 4.
4. An isolated DNA molecule according to claim 1, wherein said DNA molecule is a DNA molecule which hybridizes to a DNA molecule comprising a nucleotide sequence of SEQ. ID. Nos. 1 or 3, under stringent conditions.
5. An isolated DNA molecule according to claim 1, wherein said DNA molecule is a DNA molecule complementary to DNA molecules (a), (b), and (c).
6. An expression vector transformed with the DNA molecule of claim 1.
7. An expression vector according to claim 6, wherein the DNA molecule is in proper sense orientation and correct reading frame.
8. A host cell transformed with the DNA molecule of claim 1.

- 60 -

9. A host cell according to claim 8, wherein the host cell is selected from the group consisting of a plant cell or a bacterial cell.

10. A host cell according to claim 8, wherein the DNA molecule is transformed with an expression vector.

11. A transgenic plant transformed with the DNA molecule of claim 1.

12. A transgenic plant according to claim 11, wherein the plant is selected from the group consisting of alfalfa, rice, wheat, barley, rye, cotton, sunflower, peanut, corn, potato, sweet potato, bean, pea, chicory, lettuce, endive, cabbage, brussel sprout, beet, parsnip, turnip, cauliflower, broccoli, turnip, radish, spinach, onion, garlic, eggplant, pepper, celery, carrot, squash, pumpkin, zucchini, cucumber, apple, pear, melon, citrus, strawberry, grape, raspberry, pineapple, soybean, tobacco, tomato, sorghum, and sugarcane.

13. A transgenic plant according to claim 11, wherein the plant is selected from the group consisting of *Arabidopsis thaliana*, *Saintpaulia*, petunia, pelargonium, poinsettia, chrysanthemum, carnation, and zinnia.

14. A transgenic plant seed transformed with the DNA molecule of claim 1.

15. A transgenic plant seed according to claim 14, wherein the plant seed is selected from the group consisting of alfalfa, rice, wheat, barley, rye, cotton, sunflower, peanut, corn, potato, sweet potato, bean, pea, chicory, lettuce, endive, cabbage, brussel sprout, beet, parsnip, turnip, cauliflower, broccoli, turnip, radish, spinach, onion, garlic, eggplant, pepper, celery, carrot, squash, pumpkin, zucchini, cucumber, apple, pear, melon, citrus, strawberry, grape, raspberry, pineapple, soybean, tobacco, tomato, sorghum, and sugarcane.

16. A transgenic plant seed according to claim 14, wherein the plant seed is selected from the group consisting of *Arabidopsis thaliana*, *Saintpaulia*, petunia, pelargonium, poinsettia, chrysanthemum, carnation, and zinnia.

- 61 -

17. An isolated hypersensitive response eliciting protein or polypeptide selected from the group consisting of a protein or polypeptide having an amino acid comprising SEQ. ID. Nos. 2 or 4, and an amino acid encoded by a nucleic acid which hybridizes to a DNA molecule comprising a nucleotide sequence of SEQ. ID. Nos. 1 or 3.

18. An isolated protein or polypeptide according to claim 17, wherein the protein or polypeptide has an amino acid comprising SEQ. ID. Nos. 2 or 4.

19. An isolated protein or polypeptide according to claim 17, wherein the protein or polypeptide is encoded by a nucleic acid which hybridizes to a DNA molecule comprising a nucleotide sequence of SEQ. ID. Nos. 1 or 3.

20. A method of imparting disease resistance to plants comprising: applying a protein or polypeptide according claim 17 in a non-infectious form to a plant or plant seed under conditions effective to impart disease resistance.

21. A method according to claim 20, wherein plants are treated during said applying.

22. A method according to claim 20, wherein plant seeds are treated during said applying, said method further comprising: planting the seeds treated with the hypersensitive response elicitor in natural or artificial soil and propagating plants from the seeds planted in the soil.

23. A method of enhancing plant growth comprising: applying a protein or polypeptide according claim 17 in a non-infectious form to a plant or plant seed under conditions effective to enhance plant growth.

- 62 -

24. A method according to claim 23, wherein plants are treated during said applying.

25. A method according to claim 23, wherein plant seeds are
5 treated during said applying, said method further comprising:
planting the seeds treated with the hypersensitive response
elicitor in natural or artificial soil and
propagating plants from the seeds planted in the soil.

10 26. A method of insect control for plants comprising:
applying a protein or polypeptide according claim 17 in a non-
infectious form to a plant or plant seed under conditions effective to control insects.

15 27. A method according to claim 26, wherein plants are treated
during said applying.

28. A method according to claim 26, wherein plant seeds are
treated during said applying, said method further comprising:
planting the seeds treated with the hypersensitive response
20 elicitor in natural or artificial soil and
propagating plants from the seeds planted in the soil.

29. A method of imparting disease resistance to plants comprising:
providing a transgenic plant or plant seed transformed with a
25 DNA molecule according to claim 1 and
growing the transgenic plant or transgenic plants produced
from the transgenic plant seeds under conditions effective to impart disease resistance.

30. A method according to claim 29, wherein a transgenic plant is
30 provided.

31. A method according to claim 29, wherein a transgenic plant
seed is provided.

- 63 -

32. A method of enhancing plant growth comprising:
providing a transgenic plant or plant seed transformed with a
DNA molecule according to claim 1 and
growing the transgenic plant or transgenic plants produced
5 from the transgenic plant seeds under conditions effective to enhance plant growth.
33. A method according to claim 32, wherein a transgenic plant is
provided.
- 10 34. A method according to claim 32, wherein a transgenic plant
seed is provided.
35. A method of insect control for plants comprising:
providing a transgenic plant or plant seed transformed with a
15 DNA molecule according to claim 1 and
growing the transgenic plant or transgenic plants produced
from the transgenic plant seeds under conditions effective to control insects.
36. A method according to claim 35, wherein a transgenic plant is
20 provided.
37. A method according to claim 35, wherein a transgenic plant
seed is provided.
- 25 38. A composition comprising:
a protein or polypeptide according to claim 17 and
a carrier.
39. A composition according to claim 38 further comprising an
30 additive selected from the group consisting of fertilizer, insecticide, fungicide,
nematocide, and mixtures thereof.
40. An antibody or binding portion thereof which recognizes a protein
or polypeptide according to claim 17.

35

- 64 -

41. An antibody or binding portion thereof according to claim 40,
wherein the antibody is a monoclonal antibody.

42. An antibody or binding portion thereof according to claim 40,
5 wherein the antibody is a polyclonal antibody.

43. A method to alter the disease or hypersensitive response in a
plant comprising:
providing the plant with an antibody or binding portion thereof
10 according to claim 40 and
causing the antibody or binding portion thereof to bind to a
hypersensitive response elicitor protein or polypeptide under conditions effective to
alter disease or hypersensitive response.

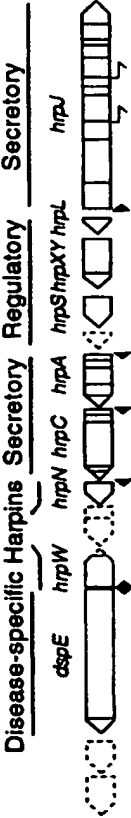


FIG. 1A

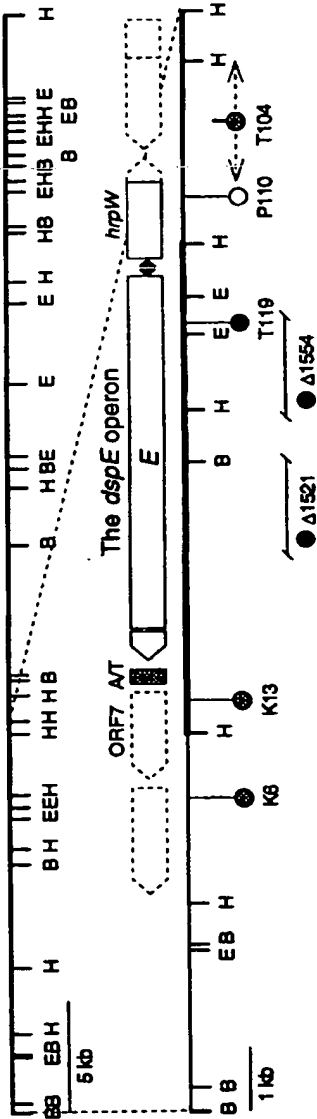


FIG. 1B

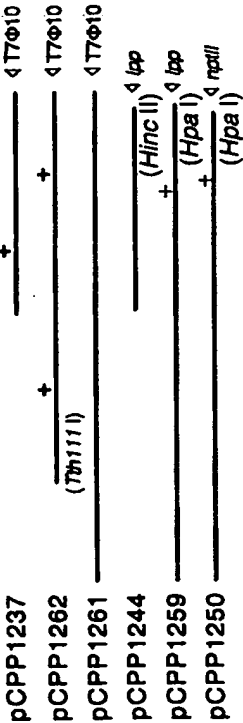


FIG. 1C

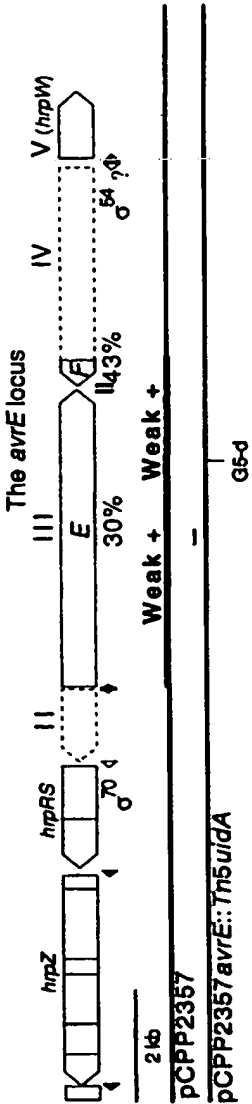
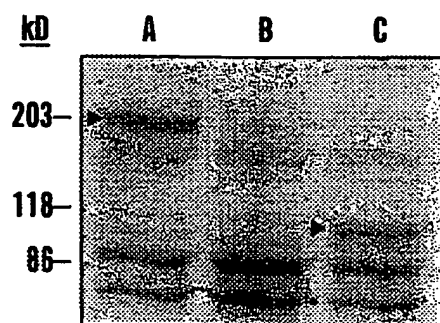


FIG. 1D

2/5

**FIG. 2**

3/5

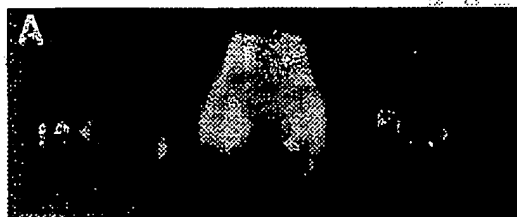


FIG. 3A

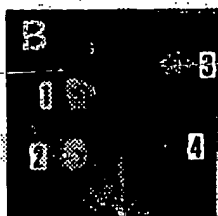


FIG. 3B

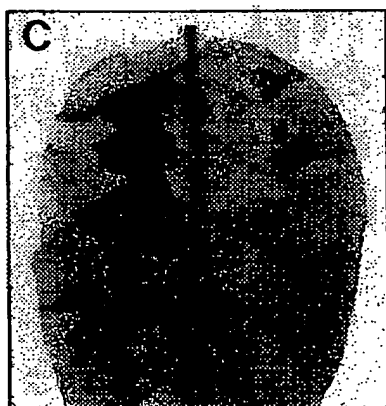


FIG. 3C

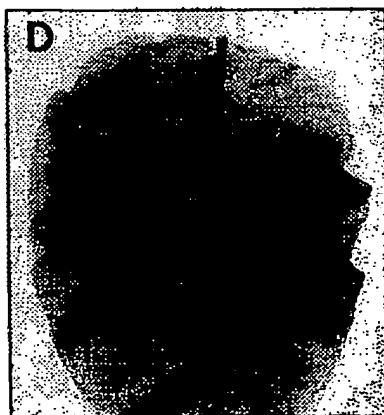
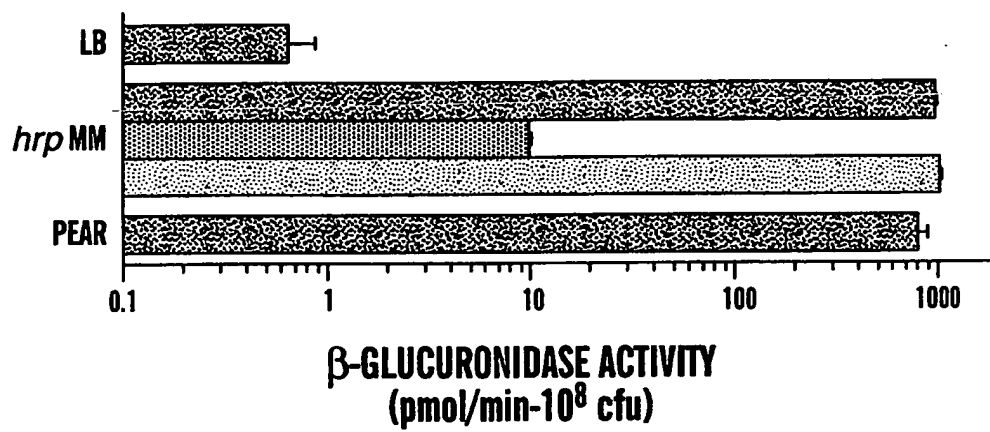


FIG. 3D

4/5

**FIG. 4**

5/5

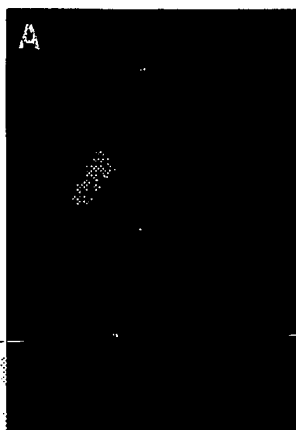


FIG. 5A

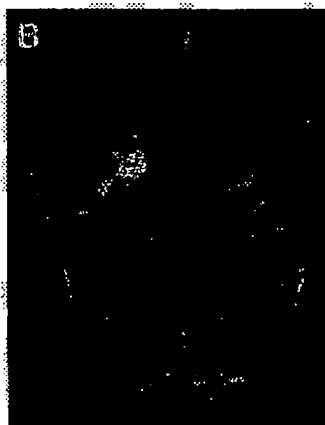


FIG. 5B

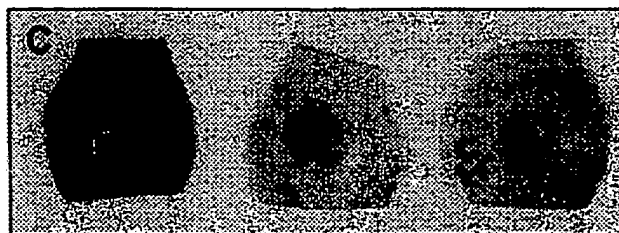


FIG. 5C

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US98/15426

A. CLASSIFICATION OF SUBJECT MATTER

IPC(6) : A01G 13/00; A61K 35/66; C12N 1/20; C12R 1/18

US CL : 530/350; 536/23.7; 435/874; 800/301

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

U.S. : 530/350; 536/23.7; 435/874; 800/301

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)


BIOSIS, MEDLINE, AGRICOLA, CAPLUS

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	WEI et al. Harpin, Elicitor of the Hypersensitive Response Produced by the Plant Pathogen Erwinia Amylovora. Science. 1992, Vol. 257, pages 85-88, see entire document.	1-16, 29-37
Y	BURR et al. Increased Potato Yields by Treatment of Seedpiece with Specific Strains of Pseudomonas Fluorescens and P. Putida. Phytopathology. 1978, Vol. 68, pages 1377-1383, see entire document.	1-16, 29-37

☒ Further documents are listed in the continuation of Box C. ☐ See patent family annex.

* Special categories of cited documents:	*T* later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
A document defining the general state of the art which is not considered to be of particular relevance	*X* document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
B earlier document published on or after the international filing date	*Y* document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art
L document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)	*Z* document member of the same patent family
O document referring to an oral disclosure, use, exhibition or other means	
P document published prior to the international filing date but later than the priority date claimed	

Date of the actual completion of the international search 24 OCTOBER 1998	Date of mailing of the international search report 03 DEC 1998
Name and mailing address of the ISA/US Commissioner of Patents and Trademarks Box PCT Washington, D.C. 20231 Facsimile No. (703) 305-3230	Authorized officer  OUSAMA M-FAIZ ZAGHMOUT Telephone No. (703) 308-0196

INTERNATIONAL SEARCH REPORT

International application No.

PCT/US98/15426

C (Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	BAILLIEUL et al. A New Elicitor of the Hypersensitive Response in Tobacco: A Fungal Glycoprotein Elicits Cell Death, Expression of Defence Genes, Production of Salicylic Acid, and Induction of Systemic Acquired Resistance. The Plant Journal. 1995, Vol. 8, No. 4, pages 551-560, see entire document.	1-16,29-37

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US98/15426

Box I Observations where certain claims were found unsearchable (Continuation of Item 1 of first sheet)

This international report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. ☐ Claims Nos.:
because they relate to subject matter not required to be searched by this Authority, namely:

2. ☐ Claims Nos.:
because they relate to parts of the international application that do not comply with the prescribed requirements to such an extent that no meaningful international search can be carried out, specifically:

3. ☐ Claims Nos.:
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

Box II Observations where unity of invention is lacking (Continuation of Item 2 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:

Please See Extra Sheet.

1. ☐ As all required additional search fees were timely paid by the applicant, this international search report covers all searchable claims.
2. ☐ As all searchable claims could be searched without effort justifying an additional fee, this Authority did not invite payment of any additional fee.
3. ☐ As only some of the required additional search fees were timely paid by the applicant, this international search report covers only those claims for which fees were paid, specifically claims Nos.:

4. ☒ No required additional search fees were timely paid by the applicant. Consequently, this international search report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:
1-16,29-37

Remark on Protest

- ☐ The additional search fees were accompanied by the applicant's protest.
☐ No protest accompanied the payment of additional search fees.

PCTWORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification ⁶ : A01G 1/00, A01H 1/00, 1/02, 5/00, 7/00, 9/00, 11/00, 13/00, 15/00, 17/00, A01N 37/18, A61K 35/78, 35/80, 38/00, C07K 1/00, 14/00, 16/00, 17/00	A1	(11) International Publication Number: WO 98/24297 (43) International Publication Date: 11 June 1998 (11.06.98)
(21) International Application Number: PCT/US97/22629 (22) International Filing Date: 4 December 1997 (04.12.97) (30) Priority Data: 60/033,230 5 December 1996 (05.12.96) US (71) Applicant: CORNELL RESEARCH FOUNDATION, INC. {US/US}; Suite 105, 20-Thornwood Drive, Ithaca, NY 14850 (US). (72) Inventors: QIU, Dewen; Apartment #2-I, 110 Dryden Road, Ithaca, NY 14850 (US). WEI, Zhong-Min; 8230 125th Court, Kirkland, WA 98034 (US). BEER, Steven, V.; 211 Hudson Street, Ithaca, NY 14850 (US). (74) Agents: GOLDMAN, Michael, L. et al.; Nixon, Hargrave, Devans & Doyle LLP, Clinton Square, P.O. Box 1051, Rochester, NY 14603 (US).	(81) Designated States: AL, AM, AT, AU, AZ, BB, BG, BR, BY, CA, CH, CN, CZ, DE, DK, EE, ES, FI, GB, GE, HU, IL, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, TJ, TM, TR, TT, UA, UG, UZ, VN, YU, ARIPO patent (GH, KE, LS, MW, SD, SZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, ML, MR, NE, SN, TD, TG). Published <i>With international search report.</i> <i>Before the expiration of the time limit for amending the</i> <i>claims and to be republished in the event of the receipt of</i> <i>amendments.</i>	
(54) Title: HYPERSENSITIVE RESPONSE INDUCED RESISTANCE IN PLANTS BY SEED TREATMENT (57) Abstract The present invention relates to a method of imparting pathogen resistance to plants. This involves applying a hypersensitive response elicitor polypeptide or protein in a non-infectious form to a plant seed under conditions where the polypeptide or protein contacts cells of the plant seed. The present invention is also directed to a pathogen resistance imparting plant seed. Alternatively, transgenic plant seeds containing a DNA molecule encoding a hypersensitive response elicitor polypeptide or protein can be planted in soil and a plant can be propagated from the planted seed under conditions effective to impart pathogen resistance to the plant.		

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	ES	Spain	LS	Lesotho	SI	Slovenia
AM	Armenia	FI	Finland	LT	Lithuania	SK	Slovakia
AT	Austria	FR	France	LU	Luxembourg	SN	Senegal
AU	Australia	GA	Gabon	LV	Latvia	SZ	Swaziland
AZ	Azerbaijan	GB	United Kingdom	MC	Monaco	TD	Chad
BA	Bosnia and Herzegovina	GE	Georgia	MD	Republic of Moldova	TG	Togo
BB	Barbados	GH	Ghana	MG	Madagascar	TJ	Tajikistan
BE	Belgium	GN	Guinea	MK	The former Yugoslav Republic of Macedonia	TM	Turkmenistan
BF	Burkina Faso	GR	Greece	ML	Mali	TR	Turkey
BG	Bulgaria	HU	Hungary	MN	Mongolia	TT	Trinidad and Tobago
BJ	Benin	IE	Ireland	MR	Mauritania	UA	Ukraine
BR	Brazil	IL	Israel	MW	Malawi	UG	Uganda
BY	Belarus	IS	Iceland	MX	Mexico	US	United States of America
CA	Canada	IT	Italy	NE	Niger	UZ	Uzbekistan
CF	Central African Republic	JP	Japan	NL	Netherlands	VN	Viet Nam
CG	Congo	KE	Kenya	NO	Norway	YU	Yugoslavia
CH	Switzerland	KG	Kyrgyzstan	NZ	New Zealand	ZW	Zimbabwe
CI	Côte d'Ivoire	KP	Democratic People's Republic of Korea	PL	Poland		
CM	Cameroon	KR	Republic of Korea	PT	Portugal		
CN	China	KZ	Kazakhstan	RO	Romania		
CU	Cuba	LC	Saint Lucia	RU	Russian Federation		
CZ	Czech Republic	LI	Liechtenstein	SD	Sudan		
DE	Germany	LK	Sri Lanka	SE	Sweden		
DK	Denmark	LR	Liberia	SG	Singapore		
EE	Estonia						

**HYPERSENSITIVE RESPONSE INDUCED RESISTANCE
IN PLANTS BY SEED TREATMENT**

5 This application claims the benefit of U.S.
Provisional Patent Application Serial No. 60/033,230,
filed December 5, 1996.

 This invention was made with support from the
U.S. Government under USDA NRI Competitive Research Grant
10 No. 91-37303-6430.

FIELD OF THE INVENTION

 The present invention relates to imparting
15 hypersensitive response induced resistance to plants by
treatment of seeds.

BACKGROUND OF THE INVENTION

20 Living organisms have evolved a complex array
of biochemical pathways that enable them to recognize and
respond to signals from the environment. These pathways
include receptor organs, hormones, second messengers, and
enzymatic modifications. At present, little is known
25 about the signal transduction pathways that are activated
during a plant's response to attack by a pathogen,
although this knowledge is central to an understanding of
disease susceptibility and resistance. A common form of
plant resistance is the restriction of pathogen
30 proliferation to a small zone surrounding the site of
infection. In many cases, this restriction is
accompanied by localized death (i.e., necrosis) of host
tissues. Together, pathogen restriction and local tissue
necrosis characterize the hypersensitive response. In
35 addition to local defense responses, many plants respond
to infection by activating defenses in uninfected parts
of the plant. As a result, the entire plant is more
resistant to a secondary infection. This systemic

- 2 -

acquired resistance can persist for several weeks or more (R.E.F. Matthews, Plant Virology (Academic Press, New York, ed. 2, 1981)) and often confers cross-resistance to unrelated pathogens (J. Kuc, in Innovative Approaches to Plant Disease Control, I. Chet, Ed. (Wiley, New York, 1987), pp. 255-274, which is hereby incorporated by reference). See also Kessman, et al., "Induction of Systemic Acquired Disease Resistance in Plants By Chemicals," Ann. Rev. Phytopathol. 32:439-59 (1994), Ryals, et al., "Systemic Acquired Resistance," The Plant Cell 8:1809-19 (Oct. 1996), and Neuenschwander, et al., "Systemic Acquired Resistance," Plant-Microbe Interactions vol. 1, G. Stacey, et al. ed. pp. 81-106 (1996), which are hereby incorporated by reference.

Expression of systemic acquired resistance is associated with the failure of normally virulent pathogens to ingress the immunized tissue (Kuc, J., "Induced Immunity to Plant Disease," Bioscience, 32:854-856 (1982), which is hereby incorporated by reference).

Establishment of systemic acquired resistance is correlated with systemic increases in cell wall hydroxyproline levels and peroxidase activity (Smith, J.A., et al., "Comparative Study of Acidic Peroxidases Associated with Induced Resistance in Cucumber, Muskmelon and Watermelon," Physiol. Mol. Plant Pathol. 14:329-338 (1988), which is hereby incorporated by reference) and with the expression of a set of nine families of so-called systemic acquired resistance gene (Ward, E.R., et al., "Coordinate Gene Activity in Response to Agents that Induce Systemic Acquired Resistance," Plant Cell 3:49-59 (1991), which is hereby incorporated by reference). Five of these defense gene families encode pathogenesis-related proteins whose physiological functions have not been established. However, some of these proteins have antifungal activity *in vitro* (Bol,

- 3 -

J.F., et al., "Plant Pathogenesis-Related Proteins Induced by Virus Infection," Ann. Rev. Phytopathol. 28:113-38 (1990), which is hereby incorporated by reference) and the constitutive expression of a bean chitinase gene in transgenic tobacco protects against infection by the fungus *Rhizoctonia solani* (Broglie, K., et al., "Transgenic Plants with Enhanced Resistance to the Fungal Pathogen *Rhizoctonia Solani*," Science 254:1194-1197 (1991), which is hereby incorporated by reference), suggesting that these systemic acquired resistance proteins may contribute to the immunized state (Uknes, S., et al., "Acquired Resistance in *Arabidopsis*," Plant Cell 4:645-656 (1992), which is hereby incorporated by reference).

Salicylic acid appears to play a signal function in the induction of systemic acquired resistance since endogenous levels increase after immunization (Malamy, J., et al., "Salicylic Acid: A Likely Endogenous Signal in the Resistance Response of Tobacco to Viral Infection," Science 250:1002-1004 (1990), which is hereby incorporated by reference) and exogenous salicylate induces systemic acquired resistance genes (Yalpani, N., et al., "Salicylic Acid is a Systemic Signal and an Inducer of Pathogenesis-Related Proteins in Virus-Infected Tobacco," Plant Cell 3:809-818 (1991), which is hereby incorporated by reference), and acquired resistance (Uknes, S., et al., "Acquired Resistance in *Arabidopsis*," Plant Cell 4:645-656 (1992), which is hereby incorporated by reference). Moreover, transgenic tobacco plants in which salicylate is destroyed by the action of a bacterial transgene encoding salicylate hydroxylase do not exhibit systemic acquired resistance (Gaffney, T., et al., "Requirement of Salicylic Acid for the Induction of Systemic Acquired Resistance," Science 261:754-56 (1993), which is hereby incorporated by

- 4 -

- reference). However, this effect may reflect inhibition of a local rather than a systemic signal function, and detailed kinetic analysis of signal transmission in cucumber suggests that salicylate may not be essential for long-distance signaling (Rasmussen, J.B., et al., "Systemic Induction of Salicylic Acid Accumulation in Cucumber after Inoculation with *Pseudomonas Syringae* pv. *Syringae*," Plant Physiol. 97:1342-1347) (1991), which is hereby incorporated by reference).
- Immunization using biotic agents has been extensively studied. Green beans were systemically immunized against disease caused by cultivar-pathogenic races of *Colletotrichum lindemuthianum* by prior infection with either cultivar-nonpathogenic races (Rahe, J.E., "Induced Resistance in *Phaseolus Vulgaris* to Bean Anthracnose," Phytopathology 59:1641-5 (1969); Elliston, J., et al., "Induced Resistance to Anthracnose at a Distance from the Site of the Inducing Interaction," Phytopathology 61:1110-12 (1971); Skipp, R., et al., "Studies on Cross Protection in the Anthracnose Disease of Bean," Physiological Plant Pathology 3:299-313 (1973), which are hereby incorporated by reference), cultivar-pathogenic races attenuated by heat in host tissue prior to symptom appearance (Rahe, J.E., et al., "Metabolic Nature of the Infection-Limiting Effect of Heat on Bean Anthracnose," Phytopathology 60:1005-9 (1970), which is hereby incorporated by reference) or nonpathogens of bean. The anthracnose pathogen of cucumber, *Colletotrichum lagenarium*, was equally effective as non-pathogenic races as an inducer of systemic protection against all races of bean anthracnose. Protection was induced by *C. lagenarium* in cultivars resistant to one or more races of *C. lindemuthianum* as well as in cultivars susceptible to all reported races of the fungus and which accordingly had

- 5 -

been referred to as 'lacking genetic resistance' to the pathogen (Elliston, J., et al., "Protection of Bean Against Anthracnose by *Colletotrichum* Species Nonpathogenic on Bean," Phytopathologische Zeitschrift 86:117-26 (1976); Elliston, J., et al., "A Comparative Study on the Development of Compatible, Incompatible and Induced Incompatible Interactions Between *Collectotrichum* Species and *Phaseolus Vulgaris*," Phytopathologische Zeitschrift 87:289-303 (1976), which are hereby incorporated by reference). These results suggest that the same mechanisms may be induced in cultivars reported as 'possessing' or 'lacking' resistance genes (Elliston, J., et al., "Relation of Phytoalexin Accumulation to Local and Systemic Protection of Bean Against Anthracnose," Phytopathologische Zeitschrift 88:114-30 (1977), which is hereby incorporated by reference). It also is apparent that cultivars susceptible to all races of *C. lindemuthianum* do not lack genes for induction of resistance mechanisms against the pathogen.

Kuc, J., et al., "Protection of Cucumber Against *Collectotrichum Lagenarium* by *Colletotrichum Lagenarium*," Physiological Plant Pathology 7:195-9 (1975), which is hereby incorporated by reference), showed that cucumber plants could be systemically protected against disease caused by *Colletotrichum lagenarium* by prior inoculation of the cotyledons or the first true leaf with the same fungus. Subsequently, cucumbers have been systemically protected against fungal, bacterial, and viral diseases by prior localized infection with either fungi, bacteria, or viruses (Hammerschmidt, R., et al., "Protection of Cucumbers Against *Colletotrichum Lagenarium* and *Cladosporium Cucumerinum*," Phytopathology 66:790-3 (1976); Jenns, A. E., et al., "Localized Infection with Tobacco Necrosis Virus Protects Cucumber Against *Colletotrichum*

- 6 -

Lagenarium," Physiological Plant Pathology 11:207-12 (1977); Caruso, F.L., et al. "Induced Resistance of Cucumber to Anthracnose and Angular Leaf Spot by *Pseudomonas Lachrymans* and *Colletotrichum Lagenarium*," Physiological Plant Pathology 14:191-201 (1979); Staub, T., et al., "Systemic Protection of Cucumber Plants Against Disease Caused by *Cladosporium Cucumerinum* and *Colletotrichum Lagenarium* by Prior Localized Infection with Either Fungus," Physiological Plant Pathology, 17:389-93 (1980); Bergstrom, G.C., et al., "Effects of Local Infection of Cucumber by *Colletotrichum Lagenarium*, *Pseudomonas Lachrymans* or Tobacco Necrosis Virus on Systemic Resistance to Cucumber Mosaic Virus," Phytopathology 72:922-6 (1982); Gessler, C., et al., "Induction of Resistance to *Fusarium* Wilt in Cucumber by Root and Foliar Pathogens," Phytopathology 72:1439-41 (1982); Basham, B., et al., "Tobacco Necrosis Virus Induces Systemic Resistance in Cucumbers Against *Sphaerotheca Fuliginea*," Physiological Plant Pathology 23:137-44 (1983), which are hereby incorporated by reference). Non-specific protection induced by infection with *C. lagenarium* or tobacco necrosis virus was effective against at least 13 pathogens, including obligatory and facultative parasitic fungi, local lesion and systemic viruses, wilt fungi, and bacteria. Similarly, protection was induced by and was also effective against root pathogens. Other curcubits, including watermelon and muskmelon have been systemically protected against *C. lagenarium* (Caruso, F.L., et al., "Protection of Watermelon and Muskmelon Against *Colletotrichum Lagenarium* by *Colletotrichum Lagenarium*," Phytopathology 67:1285-9 (1977), which is hereby incorporated by reference).

Systemic protection in tobacco has also been induced against a wide variety of diseases (Kuc, J., et

- 7 -

al., "Immunization for Disease Resistance in Tobacco,"
Recent Advances in Tobacco Science 9:179-213 (1983),
which is hereby incorporated by reference). Necrotic
lesions caused by tobacco mosaic virus enhanced
5 resistance in the upper leaves to disease caused by the
virus (Ross, A.F., et al., "Systemic Acquired Resistance
Induced by Localized Virus Infections in Plants,"
Virology 14:340-58 (1961); Ross, A.F., et al., "Systemic
Effects of Local Lesion Formation," In: Viruses of Plants
10 pp. 127-50 (1966), which are hereby incorporated by
reference). *Phytophthora parasitica* var. *nicotianae*, *P.*
tabacina and *Pseudomonas tabaci* and reduced reproduction
of the aphid *Myzus persicae* (McIntyre, J.L., et al.,
"Induction of Localized and Systemic Protection Against
15 *Phytophthora Parasitica* var. *nicotianae* by Tobacco Mosaic
Virus Infection of Tobacco Hypersensitive to the Virus,"
Physiological Plant Pathology 15:321-30 (1979); McIntyre,
J.L., et al., "Effects of Localized Infections of
Nicotiana Tabacum by Tobacco Mosaic Virus on Systemic
20 Resistance Against Diverse Pathogens and an Insect,"
Phytopathology 71:297-301 (1981), which are hereby
incorporated by reference). Infiltration of heat-killed
Pseudomonas tabacin (Lovrekovich, L., et al., "Induced
Reaction Against Wildfire Disease in Tobacco Leaves
25 Treated with Heat-Killed Bacteria," Nature 205:823-4
(1965), which is hereby incorporated by reference), and
Pseudomonas solanacearum (Sequeira, L., et al.,
"Interaction of Bacteria and Host Cell Walls: Its
Relation to Mechanisms of Induced Resistance,"
30 Physiological Plant Pathology 10:43-50 (1977), which is
hereby incorporated by reference), into tobacco leaves
induced resistance against the same bacteria used for
infiltration. Tobacco plants were also protected by the
nematode *Pratylenchus penetrans* against *P. parasitica*
35 var. *nicotiana* (McIntyre, J.L., et al. "Protection of

- 8 -

Tobacco Against *Phytophthora Parasitica* Var. *Nicotianae* by Cultivar-Nonpathogenic Races, Cell-Free Sonicates and *Pratylenchus Penetrans*," Phytopathology 68:235-9 (1978), which is hereby incorporated by reference).

- 5 Cruikshank, I.A.M., et al., "The Effect of Stem Infestation of Tobacco with *Peronospora Tabacina* Adam on Foliage Reaction to Blue Mould," Journal of the Australian Institute of Agricultural Science 26:369-72 (1960), which is hereby incorporated by reference, were
10 the first to report immunization of tobacco foliage against blue mould (i.e., *P. tabacina*) by stem injection with the fungus, which also resulted in dwarfing and premature senescence. It was recently discovered that injection external to the xylem not only alleviated
15 stunting but also promoted growth and development. Immunized tobacco plants, in both glasshouse and field experiments, were approximately 40% taller, had a 40% increase in dry weight, a 30% increase in fresh weight, and 4-6 more leaves than control plants (Tuzun, S., et
20 al., "The Effect of Stem Injections with *Peronospora Tabacina* and Metalaxyl Treatment on Growth of Tobacco and Protection Against Blue Mould in the Field," Phytopathology 74:804 (1984), which is hereby incorporated by reference). These plants flowered
25 approximately 2-3 weeks earlier than control plants (Tuzun, S., et al., "Movement of a Factor in Tobacco Infected with *Peronospora Tabacina* Adam which Systemically Protects Against Blue Mould," Physiological Plant Pathology 26:321-30 (1985), which is hereby
30 incorporated by reference).

Systemic protection does not confer absolute immunity against infection, but reduces the severity of the disease and delays symptom development. Lesion number, lesion size, and extent of sporulation of fungal

- 9 -

pathogens are all decreased. The diseased area may be reduced by more than 90%.

When cucumbers were given a 'booster' inoculation 3-6 weeks after the initial inoculation, immunization induced by *C. lagenarium* lasted through flowering and fruiting (Kuc, J., et al., "Aspects of the Protection of Cucumber Against *Colletotrichum Lagenarium* by *Colletotrichum Lagenarium*," Phytopathology 67:533-6 (1977), which is hereby incorporated by reference). Protection could not be induced once plants had set fruit. Tobacco plants were immunized for the growing season by stem injection with sporangia of *P. tabacina*. However, to prevent systemic blue mould development, this technique was only effective when the plants were above 20 cm in height.

Removal of the inducer leaf from immunized cucumber plants did not reduce the level of immunization of pre-existing expanded leaves. However, leaves which subsequently emerged from the apical bud were progressively less protected than their predecessors (Dean, R.A., et al., "Induced Systemic Protection in Cucumber: Time of Production and Movement of the 'Signal'," Phytopathology 76:966-70 (1986), which is hereby incorporated by reference). Similar results were reported by Ross, A.F., "Systemic Effects of Local Lesion Formation," In: Viruses of Plants pp. 127-50 (1966), which is hereby incorporated by reference, with tobacco (local lesion host) immunized against tobacco mosaic virus by prior infection with tobacco mosaic virus. In contrast, new leaves which emerged from scions excised from tobacco plants immunized by stem-injection with *P. tabacina* were highly protected (Tuzun, S., et al., "Transfer of Induced Resistance in Tobacco to Blue Mould (*Peronospora tabacina* Adam.) Via Callus," Phytopathology 75:1304 (1985), which is hereby incorporated by

- 10 -

reference). Plants regenerated via tissue culture from leaves of immunized plants showed a significant reduction in blue mould compared to plants regenerated from leaves of non-immunized parents. Young regenerants only showed reduced sporulation. As plants aged, both lesion development and sporulation were reduced. Other investigators, however, did not reach the same conclusion, although a significant reduction in sporulation in one experiment was reported (Lucas, J.A., et al., "Nontransmissibility to Regenerants from Protected Tobacco Explants of Induced Resistance to *Peronospora Hyoscyami*," Phytopathology 75:1222-5 (1985), which is hereby incorporated by reference).

Protection of cucumber and watermelon is effective in the glasshouse and in the field (Caruso, F.L., et al., "Field Protection of Cucumber Against *Colletotrichum Lagenarium* by *C. Lagenarium*," Phytopathology 67:1290-2 (1977), which is hereby incorporated by reference). In one trial, the total lesion area of *C. lagenarium* on protected cucumber was less than 2% of the lesion areas on unprotected control plants. Similarly, only 1 of 66 protected, challenged plants died, whereas 47 of 69 unprotected, challenged watermelons died. In extensive field trials in Kentucky and Puerto Rico, stem injection of tobacco with sporangia of *P. tabacina* was at least as effective in controlling blue mould as the best fungicide, metalaxyl. Plants were protected, leading to a yield increase of 10-25% in cured tobacco.

Induced resistance against bacteria and viruses appears to be expressed as suppression of disease symptoms or pathogen multiplication or both (Caruso, F.L., et al., "Induced Resistance of Cucumber to Anthracnose and Angular Leaf Spot by *Pseudomonas Lachrymans* and *Colletotrichum Lagenarium*," Physiological

- 11 -

Plant Pathology 14:191-201 (1979); Doss, M., et al.,
"Systemic Acquired Resistance of Cucumber to *Pseudomonas*
Lachrymans as Expressed in Suppression of Symptoms, but
not in Multiplication of Bacteria," Acta Phytopathologia
5 Academiae Scientiarum Hungaricae 16:(3-4), 269-72 (1981);
Jenns, A.E., et al., "Non-Specific Resistance to
Pathogens Induced Systemically by Local Infection of
Cucumber with Tobacco Necrosis Virus, *Colletotrichum*
Lagenarium or *Pseudomonas Lachrymans*," Phytopathologia
10 Mediterranea 18:129-34 (1979), which are hereby
incorporated by reference).

As described above, research concerning
systemic acquired resistance involves infecting plants
with infectious pathogens. Although studies in this area
15 are useful in understanding how systemic acquired
resistance works, eliciting such resistance with
infectious agents is not commercially useful, because
such plant-pathogen contact can weaken or kill plants.
The present invention is directed to overcoming this
20 deficiency.

SUMMARY OF THE INVENTION

The present invention relates to a method of
25 producing plant seeds which impart pathogen resistance to
plants grown from the seeds. This method involves
applying a hypersensitive response elicitor polypeptide
or protein in a non-infectious form to plant seeds under
conditions where the polypeptide or protein contacts
30 cells of the plant seeds.

As an alternative to applying a hypersensitive
response elicitor polypeptide or protein to plant seeds
in order to impart pathogen resistance to plants grown
from the seeds, transgenic seeds can be utilized. This
35 involves providing a transgenic plant seed transformed

- 12 -

with a DNA molecule encoding a hypersensitive response elicitor polypeptide or protein and planting that seed in soil. A plant is then propagated from the planted seed under conditions effective to impart pathogen resistance to the plant.

Another aspect of the present invention relates to a pathogen-resistance imparting plant seed to which a non-infectious hypersensitive response elicitor polypeptide or protein has been applied.

The present invention has the potential to: treat plant diseases which were previously untreatable; treat diseases systemically that one would not want to treat separately due to cost; and avoid the use of agents that have an unpredictable effect on the environment and even the plants. The present invention can impart resistance without using agents which are harmful to the environment or pathogenic to the plant seeds being treated or to plants situated near the location that treated seeds are planted. Since the present invention involves use of a natural product that is fully and rapidly biodegradable, the environment would not be contaminated.

DETAILED DESCRIPTION OF THE INVENTION

The present invention relates to a method of producing plant seeds which impart pathogen resistance to plants grown from the seeds. This method involves applying a hypersensitive response elicitor polypeptide or protein in a non-infectious form to a plant seed under conditions effective to impart disease resistance to a plant grown from the seed.

As an alternative to applying a hypersensitive response elicitor polypeptide or protein to plant seeds in order to impart pathogen resistance to plants grown

- 13 -

from the seeds, transgenic seeds can be utilized. This involves providing a transgenic plant seed transformed with a DNA molecule encoding a hypersensitive response elicitor polypeptide or protein and planting that seed in soil. A plant is then propagated from the planted seed under conditions effective to impart pathogen resistance to the plant.

Another aspect of the present invention relates to a pathogen-resistance imparting plant seed to which a non-infectious hypersensitive response elicitor polypeptide or protein has been applied.

The hypersensitive response elicitor polypeptide or protein utilized in the present invention can correspond to hypersensitive response elicitor polypeptides or proteins derived from a wide variety of fungal and bacterial pathogens. Such polypeptides or proteins are able to elicit local necrosis in plant tissue contacted by the elicitor.

Examples of suitable bacterial sources of polypeptide or protein elicitors include *Erwinia*, *Pseudomonas*, and *Xanthomonas* species (e.g., the following bacteria: *Erwinia amylovora*, *Erwinia chrysanthemi*, *Erwinia stewartii*, *Erwinia carotovora*, *Pseudomonas syringae*, *Pseudomonas solanacearum*, *Xanthomonas campestris*, or mixtures thereof).

An example of a fungal source of a hypersensitive response elicitor protein or polypeptide is *Phytophthora*. Suitable species of such fungal pathogens include *Phytophthora parasitica*, *Phytophthora cryptogea*, *Phytophthora cinnamomi*, *Phytophthora capsici*, *Phytophthora megasperma*, and *Phytophthora citrophthora*.

The embodiment of the present invention where the hypersensitive response elicitor polypeptide or protein is applied to the plant seed can be carried out

- 14 -

in a number of ways, including: 1) application of an isolated elicitor polypeptide or protein; 2) application of bacteria which do not cause disease and are transformed with genes encoding a hypersensitive response elicitor polypeptide or protein; and 3) application of bacteria which cause disease in some plant species (but not in those to which they are applied) and naturally contain a gene encoding the hypersensitive response elicitor polypeptide or protein. In addition, seeds in accordance with the present invention can be recovered from plants which have been treated with a hypersensitive response elicitor protein or polypeptide in accordance with the present invention.

In one embodiment of the present invention, the hypersensitive response elicitor polypeptides or proteins to be applied can be isolated from their corresponding organisms and applied to plants. Such isolation procedures are well known, as described in Arlat, M., F. Van Gijsegem, J. C. Huet, J. C. Pemollet, and C. A. Boucher, "PopA1, a Protein which Induces a Hypersensitive-like Response in Specific *Petunia* Genotypes is Secreted via the Hrp Pathway of *Pseudomonas solanacearum*," EMBO J. 13:543 - 553 (1994); He, S. Y., H. C. Huang, and A. Collmer, "*Pseudomonas syringae* pv. *syringae* Harpin_{psb}: a Protein that is Secreted via the Hrp Pathway and Elicits the Hypersensitive Response in Plants," Cell 73:1255-1266 (1993); and Wei, Z.-M., R. J. Laby, C. H. Zumoff, D. W. Bauer, S.-Y. He, A. Collmer, and S. V. Beer, "Harpin Elicitor of the Hypersensitive Response Produced by the Plant Pathogen *Erwinia amylovora*," Science 257:85-88 (1992), which are hereby incorporated by reference. See also pending U.S. Patent Application Serial Nos. 08/200,024 and 08/062,024, which are hereby incorporated by reference. Preferably, however, the isolated hypersensitive response elicitor

- 15 -

polypeptides or proteins of the present invention are produced recombinantly and purified as described below.

In other embodiments of the present invention, the hypersensitive response elicitor polypeptide or protein of the present invention can be applied to plant seeds by applying bacteria containing genes encoding the hypersensitive response elicitor polypeptide or protein. Such bacteria must be capable of secreting or exporting the polypeptide or protein so that the elicitor can contact plant seed cells. In these embodiments, the hypersensitive response elicitor polypeptide or protein is produced by the bacteria after application to the seeds or just prior to introduction of the bacteria to the seeds.

In one embodiment of the bacterial application mode of the present invention, the bacteria to be applied do not cause the disease and have been transformed (e.g., recombinantly) with genes encoding a hypersensitive response elicitor polypeptide or protein. For example, *E. coli*, which do not elicit a hypersensitive response in plants, can be transformed with genes encoding a hypersensitive response elicitor polypeptide and other related proteins required for production and secretion of the elicitor which is then applied to plant seeds. Expression of this polypeptide or protein can then be caused to occur. Bacterial species (other than *E. coli*) can also be used in this embodiment of the present invention.

In another embodiment of the bacterial application mode of the present invention, the bacteria do cause disease and naturally contain a gene encoding a hypersensitive response elicitor polypeptide or protein. Examples of such bacteria are noted above. However, in this embodiment these bacteria are applied to plant seeds for plants which are not susceptible to the disease

- 16 -

carried by the bacteria. For example, *Erwinia amylovora* causes disease in apple or pear but not in tomato. However, such bacteria will elicit a hypersensitive response in tomato. Accordingly, in accordance with this embodiment of the present invention, *Erwinia amylovora* can be applied to tomato seeds to impart pathogen resistance without causing disease in plants of that species.

The hypersensitive response elicitor polypeptide or protein from *Erwinia chrysanthemi* has an amino acid sequence corresponding to SEQ. ID. No. 1 as follows:

15	Met Gln Ile Thr Ile Lys Ala His Ile Gly Gly Asp Leu Gly Val Ser	1 5 10 15
	Gly Leu Gly Ala Gln Gly Leu Lys Gly Leu Asn Ser Ala Ala Ser Ser	20 25 30
20	Leu Gly Ser Ser Val Asp Lys Leu Ser Ser Thr Ile Asp Lys Leu Thr	35 40 45
	Ser Ala Leu Thr Ser Met Met Phe Gly Gly Ala Leu Ala Gln Gly Leu	50 55 60
25	Gly Ala Ser Ser Lys Gly Leu Gly Met Ser Asn Gln Leu Gly Gln Ser	65 70 75 80
	Phe Gly Asn Gly Ala Gln Gly Ala Ser Asn Leu Leu Ser Val Pro Lys	85 90 95
30	Ser Gly Gly Asp Ala Leu Ser Lys Met Phe Asp Lys Ala Leu Asp Asp	100 105 110
	Leu Leu Gly His Asp Thr Val Thr Lys Leu Thr Asn Gln Ser Asn Gln	115 120 125
	Leu Ala Asn Ser Met Leu Asn Ala Ser Gln Met Thr Gln Gly Asn Met	130 135 140
40	Asn Ala Phe Gly Ser Gly Val Asn Asn Ala Leu Ser Ser Ile Leu Gly	145 150 155 160
	Asn Gly Leu Gly Gln Ser Met Ser Gly Phe Ser Gln Pro Ser Leu Gly	165 170 175
45	Ala Gly Gly Leu Gln Gly Leu Ser Gly Ala Gly Ala Phe Asn Gln Leu	180 185 190
50	Gly Asn Ala Ile Gly Met Gly Val Gly Gln Asn Ala Ala Leu Ser Ala	195 200 205

- 17 -

Leu Ser Asn Val Ser Thr His Val Asp Gly Asn Asn Arg His Phe Val
 210 215 220
 5 Asp Lys Glu Asp Arg Gly Met Ala Lys Glu Ile Gly Gln Phe Met Asp
 225 230 235 240
 Gln Tyr Pro Glu Ile Phe Gly Lys Pro Glu Tyr Gln Lys Asp Gly Trp
 245 250 255
 10 Ser Ser Pro Lys Thr Asp Asp Lys Ser Trp Ala Lys Ala Leu Ser Lys
 260 265 270
 Pro Asp Asp Asp Gly Met Thr Gly Ala Ser Met Asp Lys Phe Arg Gln
 275 280 285
 15 Ala Met Gly Met Ile Lys Ser Ala Val Ala Gly Asp Thr Gly Asn Thr
 290 295 300
 Asn Leu Asn Leu Arg Gly Ala Gly Gly Ala Ser Leu Gly Ile Asp Ala
 305 310 315 320
 20 Ala Val Val Gly Asp Lys Ile Ala Asn Met Ser Leu Gly Lys Leu Ala
 325 330 335
 25 Asn Ala

This hypersensitive response elicitor polypeptide or
 protein has a molecular weight of 34 kDa, is heat stable,
 30 has a glycine content of greater than 16%, and contains
 substantially no cysteine. The *Erwinia chrysanthemi*
 hypersensitive response elicitor polypeptide or protein
 is encoded by a DNA molecule having a nucleotide sequence
 corresponding to SEQ. ID. No. 2 as follows:

35 CGATTTTACC CGGGTGAACG TGCTATGACC GACAGCATCA CGGTATTGCA CACCGTTACG 60
 GCGTTTATGG CCGCGATGAA CCGGCATCAG GCGGCGCGCT GGTCGCCGCA ATCCGGCGTC 120
 40 GATCTGGTAT TTCAGTTTGG GGACACCGGG CGTGAACTCA TGATGCAGAT TCAGCCGGGG 180
 CAGCAATATC CCGGCATGTT GCGCACGCTG CTCGCTCGTC GTTATCAGCA GGCGGCAGAG 240
 TGCGATGGCT GCCATCTGTG CCTGAACGGC AGCGATGTAT TGATCCTCTG GTGGCCGCTG 300
 45 CCGTCGGATC CCGGCAGTTA TCCGCAGGTG ATCGAACGTT TGTTTGAACT GGCGGGAATG 360
 ACGTTGCCGT CGCTATCCAT AGCACCGACG GCGCGTCCGC AGACAGGGAA CGGACGCGCC 420
 50 CGATCATTAA GATAAAGGCG GCTTTTTTTTA TTGCAAACG GTAACGGTGA GGAACCGTTT 480
 CACCGTCGGC GTCACCTCAGT AACAAATATC CATCATGATG CCTACATCGG GATCGGCGTG 540
 GGCATCCGTT GCAGATACTT TTGCGAACAC CTGACATGAA TGAGGAAACG AAATTATGCA 600
 55 AATTACGATC AAAGCGCACA TCGGCGGTGA TTTGGGCGTC TCCGGTCTGG GGCTGGGTGC 660

- 18 -

	TCAGGGACTG	AAAGGACTGA	ATTCCGCGGC	TTCATCGCTG	GGTTCCAGCG	TGGATAAACT	720
	GAGCAGCACC	ATCGATAAGT	TGACCTCCGC	GCTGACTTCG	ATGATGTTTG	GCGGCGCGCT	780
5	GGCGCAGGGG	CTGGGCGCCA	GCTCGAAGGG	GCTGGGGATG	AGCAATCAAC	TGGGCCAGTC	840
	TTTCGGCAAT	GGCGCGCAGG	GTGCGAGCAA	CCTGCTATCC	GTACCGAAAT	CCGGCGGCGA	900
	TGCGTTGTCA	AAAATGTTTG	ATAAAGCGCT	GGACGATCTG	CTGGGTCATG	ACACCGTGAC	960
10	CAAGCTGACT	AACCAGAGCA	ACCAACTGGC	TAATTCAATG	CTGAACGCCA	GCCAGATGAC	1020
	CCAGGGTAAT	ATGAATGCGT	TCGGCAGCGG	TGTGAACAAC	GCACTGTCGT	CCATTCTCGG	1080
15	CAACGGTCTC	GGCCAGTCGA	TGAGTGGCTT	CTCTCAGCCT	TCTCTGGGGG	CAGGCGGCTT	1140
	GCAGGGCCTG	AGCGGCGCGG	GTGCATTCAA	CCAGTTGGGT	AATGCCATCG	GCATGGGCGT	1200
	GGGGCAGAAT	GCTGCGCTGA	GTGCGTTGAG	TAACGTCAGC	ACCCACGTAG	ACGGTAACAA	1260
20	CCGCCACTTT	GTAGATAAAG	AAGATCGCGG	CATGGCGAAA	GAGATCGGCC	AGTTTATGGA	1320
	TCAGTATCCG	GAAATATTCG	GTAAACCGGA	ATACCAGAAA	GATGGCTGGA	GTTGCGCGAA	1380
25	GACGGACGAC	AAATCCTGGG	CTAAAGCGCT	GAGTAAACCG	GATGATGACG	GTATGACCGG	1440
	CGCCAGCATG	GACAAATTCC	GTCAGGCGAT	GGGTATGATC	AAAAGCGCGG	TGGCGGGTGA	1500
	TACCGGCAAT	ACCAACCTGA	ACCTGCGTGG	CGCGGGCGGT	GCATCGCTGG	GTATCGATGC	1560
30	GGCTGTGCTC	GGCGATAAAA	TAGCCAACAT	GTCGCTGGGT	AAGCTGGCCA	ACGCCTGATA	1620
	ATCTGTGCTG	GCCTGATAAA	GCGGAAACGA	AAAAAGAGAC	GGGGAAGCCT	GTCTCTTTTC	1680
35	TTATTATGCG	GTTTATGCGG	TTACCTGGAC	CGGTTAATCA	TCGTCATCGA	TCTGGTACAA	1740
	ACGCACATTT	TCCCGTTTCAT	TCGCGTCGTT	ACGCGCCACA	ATCGCGATGG	CATCTTCCTC	1800
	GTCGCTCAGA	TTGCGCGGCT	GATGGGGAAC	GCCGGGTGGA	ATATAGAGAA	ACTCGCCGGC	1860
40	CAGATGGAGA	CACGTCTGCG	ATAAATCTGT	GCCGTAACGT	GTTTCTATCC	GCCCCTTTAG	1920
	CAGATAGATT	GCGGTTTTCGT	AATCAACATG	GTAATGCGGT	TCCGCCTGTG	CGCCGGCCGG	1980
45	GATCACCACA	ATATTCATAG	AAAGCTGTCT	TGCACCTACC	GTATCGCGGG	AGATACCGAC	2040
	AAAATAGGGC	AGTTTTTTCG	TGGTATCCGT	GGGGTGTTCC	GGCCTGACAA	TCTTGAGTTG	2100
50	GTTTCGTCATC	ATCTTTCTCC	ATCTGGGCGA	CCTGATCGGT	T		2141

The hypersensitive response elicitor polypeptide or protein derived from *Erwinia amylovora* has an amino acid sequence corresponding to SEQ. ID. No. 3 as follows:

55

Met Ser Leu Asn Thr Ser Gly Leu Gly Ala Ser Thr Met Gln Ile Ser
 1 5 10 15

- 19 -

	Ile	Gly	Gly	Ala	Gly	Gly	Asn	Asn	Gly	Leu	Leu	Gly	Thr	Ser	Arg	Gln	
				20					25					30			
5	Asn	Ala	Gly	Leu	Gly	Gly	Asn	Ser	Ala	Leu	Gly	Leu	Gly	Gly	Gly	Asn	
			35					40					45				
	Gln	Asn	Asp	Thr	Val	Asn	Gln	Leu	Ala	Gly	Leu	Leu	Thr	Gly	Met	Met	
		50					55					60					
10	Met	Met	Met	Ser	Met	Met	Gly	Gly	Gly	Gly	Leu	Met	Gly	Gly	Gly	Leu	
	65					70					75					80	
	Gly	Gly	Gly	Leu	Gly	Asn	Gly	Leu	Gly	Gly	Ser	Gly	Gly	Leu	Gly	Glu	
					85					90					95		
15	Gly	Leu	Ser	Asn	Ala	Leu	Asn	Asp	Met	Leu	Gly	Gly	Ser	Leu	Asn	Thr	
				100					105					110			
	Leu	Gly	Ser	Lys	Gly	Gly	Asn	Asn	Thr	Thr	Ser	Thr	Thr	Asn	Ser	Pro	
20			115					120					125				
	Leu	Asp	Gln	Ala	Leu	Gly	Ile	Asn	Ser	Thr	Ser	Gln	Asn	Asp	Asp	Ser	
		130					135					140					
25	Thr	Ser	Gly	Thr	Asp	Ser	Thr	Ser	Asp	Ser	Ser	Asp	Pro	Met	Gln	Gln	
	145					150					155				160		
	Leu	Leu	Lys	Met	Phe	Ser	Glu	Ile	Met	Gln	Ser	Leu	Phe	Gly	Asp	Gly	
					165				170						175		
30	Gln	Asp	Gly	Thr	Gln	Gly	Ser	Ser	Ser	Gly	Gly	Lys	Gln	Pro	Thr	Glu	
				180					185					190			
	Gly	Glu	Gln	Asn	Ala	Tyr	Lys	Lys	Gly	Val	Thr	Asp	Ala	Leu	Ser	Gly	
35			195					200					205				
	Leu	Met	Gly	Asn	Gly	Leu	Ser	Gln	Leu	Leu	Gly	Asn	Gly	Gly	Leu	Gly	
		210					215					220					
40	Gly	Gly	Gln	Gly	Gly	Asn	Ala	Gly	Thr	Gly	Leu	Asp	Gly	Ser	Ser	Leu	
	225					230					235					240	
	Gly	Gly	Lys	Gly	Leu	Gln	Asn	Leu	Ser	Gly	Pro	Val	Asp	Tyr	Gln	Gln	
					245					250					255		
45	Leu	Gly	Asn	Ala	Val	Gly	Thr	Gly	Ile	Gly	Met	Lys	Ala	Gly	Ile	Gln	
				260					265					270			
	Ala	Leu	Asn	Asp	Ile	Gly	Thr	His	Arg	His	Ser	Ser	Thr	Arg	Ser	Phe	
50			275					280					285				
	Val	Asn	Lys	Gly	Asp	Arg	Ala	Met	Ala	Lys	Glu	Ile	Gly	Gln	Phe	Met	
		290					295					300					
55	Asp	Gln	Tyr	Pro	Glu	Val	Phe	Gly	Lys	Pro	Gln	Tyr	Gln	Lys	Gly	Pro	
	305					310					315					320	
	Gly	Gln	Glu	Val	Lys	Thr	Asp	Asp	Lys	Ser	Trp	Ala	Lys	Ala	Leu	Ser	
					325					330					335		
60	Lys	Pro	Asp	Asp	Asp	Gly	Met	Thr	Pro	Ala	Ser	Met	Glu	Gln	Phe	Asn	
			340						345					350			

- 20 -

Lys Ala Lys Gly Met Ile Lys Arg Pro Met Ala Gly Asp Thr Gly Asn
 355 360 365
 5 Gly Asn Leu Gln Ala Arg Gly Ala Gly Gly Ser Ser Leu Gly Ile Asp
 370 375 380
 Ala Met Met Ala Gly Asp Ala Ile Asn Asn Met Ala Leu Gly Lys Leu
 385 390 395 400
 10 Gly Ala Ala

This hypersensitive response elicitor polypeptide or
 15 protein has a molecular weight of about 39 kDa, it has a
 pI of approximately 4.3, and is heat stable at 100°C for
 at least 10 minutes. This hypersensitive response
 elicitor polypeptide or protein has substantially no
 cysteine. The hypersensitive response elicitor
 20 polypeptide or protein derived from *Erwinia amylovora* is
 more fully described in Wei, Z.-M., R. J. Laby, C. H.
 Zumoff, D. W. Bauer, S.-Y. He, A. Collmer, and S. V.
 Beer, "Harpin, Elicitor of the Hypersensitive Response
 Produced by the Plant Pathogen *Erwinia amylovora*,"
 25 Science 257:85-88 (1992), which is hereby incorporated by
 reference. The DNA molecule encoding this polypeptide or
 protein has a nucleotide sequence corresponding to SEQ.
 ID. No. 4 as follows:

30	AAGCTTCGGC ATGGCACGTT TGACCGTTGG GTCGGCAGGG TACGTTTGAA TTATTCATAA	60
	GAGGAATACG TTATGAGTCT GAATACAAGT GGGCTGGGAG CGTCAACGAT GCAAATTTCT	120
	ATCGGCGGTG CGGGCGGAAA TAACGGGTTG CTGGGTACCA GTCGCCAGAA TGCTGGGTTG	180
35	GGTGGCAATT CTGCACTGGG GCTGGGCGGC GGTAATCAAA ATGATACCGT CAATCAGCTG	240
	GCTGGCTTAC TCACCGGCAT GATGATGATG ATGAGCATGA TGGGCGGTGG TGGGCTGATG	300
40	GGCGGTGGCT TAGGCGGTGG CTTAGGTAAT GGCTTGGGTG GCTCAGGTGG CCTGGGCGAA	360
	GGACTGTCTGA ACGCGCTGAA CGATATGTTA GGCGGTTCGC TGAACACGCT GGGCTCGAAA	420
	GGCGGCAACA ATACCACTTC AACAACAAAT TCCCCGCTGG ACCAGGCGCT GGGTATTAAAC	480
45	TCAACGTCCC AAAACGACGA TTCCACCTCC GGCACAGATT CCACCTCAGA CTCCAGCGAC	540
	CCGATGCAGC AGCTGCTGAA GATGTTTCAGC GAGATAATGC AAAGCCTGTT TGGTGATGGG	600
50	CAAGATGGCA CCCAGGGCAG TTCCTCTGGG GGCAAGCAGC CGACCGAAGG CGAGCAGAAC	660

- 21 -

	GCCTATAAAA AAGGAGTCAC TGATGCGCTG TCGGGCCTGA TGGGTAATGG TCTGAGCCAG	720
	CTCCTTGGCA ACGGGGGA CT GGGAGGTGGT CAGGGCGGTA ATGCTGGCAC GGGTCTTGAC	780
5	GGTTCGTCGC TGGGCGGCAA AGGGCTGCAA AACCTGAGCG GGCCGGTGGA CTACCAGCAG	840
	TTAGGTAACG CCGTGGGTAC CCGTATCGGT ATGAAAGCGG GCATTTCAGGC GCTGAATGAT	900
	ATCGGTACGC ACAGGCACAG TTCAACCCGT TCTTTCGTCA ATAAAGGCGA TCGGGCGATG	960
10	GCGAAGGAAA TCGGTCAGTT CATGGACCAG TATCCTGAGG TGTTTGGCAA GCCGCAGTAC	1020
	CAGAAAGGCC CGGGTCAGGA GGTGAAAACC GATGACAAAT CATGGGCAAA AGCACTGAGC	1080
15	AAGCCAGATG ACGACGGAAT GACACCAGCC AGTATGGAGC AGTTCAACAA AGCCAAGGGC	1140
	ATGATCAAAA GGCCCATGGC GGGTGATACC GGCAACGGCA ACCTGCAGGC ACGCGGTGCC	1200
	GGTGGTTCTT CGCTGGGTAT TGATGCCATG ATGGCCGGTG ATGCCATTAA CAATATGGCA	1260
20	CTTGGCAAGC TGGGCGCGGC TTAAGCTT	1288

The hypersensitive response elicitor

25 polypeptide or protein derived from *Pseudomonas syringae* has an amino acid sequence corresponding to SEQ. ID. No. 5 as follows:

30	Met Gln Ser Leu Ser Leu Asn Ser Ser Ser Leu Gln Thr Pro Ala Met	1 5 10 15
	Ala Leu Val Leu Val Arg Pro Glu Ala Glu Thr Thr Gly Ser Thr Ser	20 25 30
35	Ser Lys Ala Leu Gln Glu Val Val Val Lys Leu Ala Glu Glu Leu Met	35 40 45
	Arg Asn Gly Gln Leu Asp Asp Ser Ser Pro Leu Gly Lys Leu Leu Ala	50 55 60
40	Lys Ser Met Ala Ala Asp Gly Lys Ala Gly Gly Gly Ile Glu Asp Val	65 70 75 80
45	Ile Ala Ala Leu Asp Lys Leu Ile His Glu Lys Leu Gly Asp Asn Phe	85 90 95
	Gly Ala Ser Ala Asp Ser Ala Ser Gly Thr Gly Gln Gln Asp Leu Met	100 105 110
50	Thr Gln Val Leu Asn Gly Leu Ala Lys Ser Met Leu Asp Asp Leu Leu	115 120 125
	Thr Lys Gln Asp Gly Gly Thr Ser Phe Ser Glu Asp Asp Met Pro Met	130 135 140
55	Leu Asn Lys Ile Ala Gln Phe Met Asp Asp Asn Pro Ala Gln Phe Pro	145 150 155 160

- 22 -

Lys Pro Asp Ser Gly Ser Trp Val Asn Glu Leu Lys Glu Asp Asn Phe
 165 170 175
 5 Leu Asp Gly Asp Glu Thr Ala Ala Phe Arg Ser Ala Leu Asp Ile Ile
 180 185 190
 Gly Gln Gln Leu Gly Asn Gln Gln Ser Asp Ala Gly Ser Leu Ala Gly
 195 200 205
 10 Thr Gly Gly Gly Leu Gly Thr Pro Ser Ser Phe Ser Asn Asn Ser Ser
 210 215 220
 Val Met Gly Asp Pro Leu Ile Asp Ala Asn Thr Gly Pro Gly Asp Ser
 225 230 235 240
 15 Gly Asn Thr Arg Gly Glu Ala Gly Gln Leu Ile Gly Glu Leu Ile Asp
 245 250 255
 Arg Gly Leu Gln Ser Val Leu Ala Gly Gly Gly Leu Gly Thr Pro Val
 260 265 270
 20 Asn Thr Pro Gln Thr Gly Thr Ser Ala Asn Gly Gly Gln Ser Ala Gln
 275 280 285
 25 Asp Leu Asp Gln Leu Leu Gly Gly Leu Leu Leu Lys Gly Leu Glu Ala
 290 295 300
 Thr Leu Lys Asp Ala Gly Gln Thr Gly Thr Asp Val Gln Ser Ser Ala
 305 310 315 320
 30 Ala Gln Ile Ala Thr Leu Leu Val Ser Thr Leu Leu Gln Gly Thr Arg
 325 330 335
 35 Asn Gln Ala Ala Ala
 340

This hypersensitive response elicitor polypeptide or
 protein has a molecular weight of 34-35 kDa. It is rich
 40 in glycine (about 13.5%) and lacks cysteine and tyrosine.
 Further information about the hypersensitive response
 elicitor derived from *Pseudomonas syringae* is found in
 He, S. Y., H. C. Huang, and A. Collmer, "*Pseudomonas*
syringae pv. *syringae* Harpin_{PS}: a Protein that is
 45 Secreted via the Hrp Pathway and Elicits the
 Hypersensitive Response in Plants," Cell 73:1255-1266
 (1993), which is hereby incorporated by reference. The
 DNA molecule encoding the hypersensitive response
 elicitor from *Pseudomonas syringae* has a nucleotide
 50 sequence corresponding to SEQ. ID. No. 6 as follows:

- 23 -

ATGCAGAGTC TCAGTCTTAA CAGCAGCTCG CTGCAAACCC CGGCAATGGC CTTGTCTCTG 60
 GTACGTCTCTG AAGCCGAGAC GACTGGCAGT ACGTCGAGCA AGGCGCTTCA GGAAGTTGTC 120
 5 GTGAAGCTGG CCGAGGAACT GATGCGCAAT GGTCAACTCG ACGACAGCTC GCCATTGGGA 180
 AAACTGTTGG CCAAGTCGAT GGCCGCAGAT GGCAAGGCGG GCGGCGGTAT TGAGGATGTC 240
 10 ATCGCTGCGC TGGACAAGCT GATCCATGAA AAGTCGGTG ACAACTTCGG CGCGTCTGCG 300
 GACAGCGCCT CGGGTACCGG ACAGCAGGAC CTGATGACTC AGGTGCTCAA TGGCCTGGCC 360
 AAGTCGATGC TCGATGATCT TCTGACCAAG CAGGATGGCG GGACAAGCTT CTCCGAAGAC 420
 15 GATATGCCGA TGCTGAACAA GATCGCGCAG TTCATGGATG ACAATCCCGC ACAGTTTCCC 480
 AAGCCGGACT CGGGCTCCTG GGTGAACGAA CTCAAGGAAG ACAACTTCCT TGATGGCGAC 540
 20 GAAACGGCTG CGTTCCGTTT GGCACCTCGAC ATCATTGGCC AGCAACTGGG TAATCAGCAG 600
 AGTGACGCTG GCAGTCTGGC AGGGACGGGT GGAGGTCTGG GCACTCCGAG CAGTTTTTCC 660
 AACAACTCGT CCGTGATGGG TGATCCGCTG ATCGACGCCA ATACCGGTCC CGGTGACAGC 720
 25 GGCAATACCC GTGGTGAAGC GGGGCAACTG ATCGGCGAGC TTATCGACCG TGGCCTGCAA 780
 TCGGTATTGG CCGGTGGTGG ACTGGGCACA CCCGTAAACA CCCCAGAC CCGTACGTCG 840
 30 GCGAATGGCG GACAGTCCGC TCAGGATCTT GATCAGTTGC TGGGCGGCTT GCTGCTCAAG 900
 GGCCTGGAGG CAACGCTCAA GGATGCCGGG CAAACAGGCA CCGACGTGCA GTCGAGCGCT 960
 GCGCAAATCG CCACCTTGCT GGTCAGTACG CTGCTGCAAG GCACCCGCAA TCAGGCTGCA 1020
 35 GCCTGA

1026

The hypersensitive response elicitor
 polypeptide or protein derived from *Pseudomonas*
 40 *solanacearum* has an amino acid sequence corresponding to
 SEQ. ID. No. 7 as follows:

Met Ser Val Gly Asn Ile Gln Ser Pro Ser Asn Leu Pro Gly Leu Gln
 1 5 10 15
 45 Asn Leu Asn Leu Asn Thr Asn Thr Asn Ser Gln Gln Ser Gly Gln Ser
 20 25 30
 Val Gln Asp Leu Ile Lys Gln Val Glu Lys Asp Ile Leu Asn Ile Ile
 35 40 45
 50 Ala Ala Leu Val Gln Lys Ala Ala Gln Ser Ala Gly Gly Asn Thr Gly
 50 55 60
 55 Asn Thr Gly Asn Ala Pro Ala Lys Asp Gly Asn Ala Asn Ala Gly Ala
 65 70 75 80
 Asn Asp Pro Ser Lys Asn Asp Pro Ser Lys Ser Gln Ala Pro Gln Ser
 85 90 95

- 24 -

Ala Asn Lys Thr Gly Asn Val Asp Asp Ala Asn Asn Gln Asp Pro Met
 100 105 110
 5 Gln Ala Leu Met Gln Leu Leu Glu Asp Leu Val Lys Leu Leu Lys Ala
 115 120 125
 Ala Leu His Met Gln Gln Pro Gly Gly Asn Asp Lys Gly Asn Gly Val
 130 135 140
 10 Gly Gly Ala Asn Gly Ala Lys Gly Ala Gly Gly Gln Gly Gly Leu Ala
 145 150 155 160
 Glu Ala Leu Gln Glu Ile Glu Gln Ile Leu Ala Gln Leu Gly Gly Gly
 165 170 175
 15 Gly Ala Gly Ala Gly Gly Ala Gly Gly Gly Val Gly Gly Ala Gly Gly
 180 185 190
 Ala Asp Gly Gly Ser Gly Ala Gly Gly Ala Gly Gly Ala Asn Gly Ala
 195 200 205
 20 Asp Gly Gly Asn Gly Val Asn Gly Asn Gln Ala Asn Gly Pro Gln Asn
 210 215 220
 25 Ala Gly Asp Val Asn Gly Ala Asn Gly Ala Asp Asp Gly Ser Glu Asp
 225 230 235 240
 Gln Gly Gly Leu Thr Gly Val Leu Gln Lys Leu Met Lys Ile Leu Asn
 245 250 255
 30 Ala Leu Val Gln Met Met Gln Gln Gly Gly Leu Gly Gly Gly Asn Gln
 260 265 270
 Ala Gln Gly Gly Ser Lys Gly Ala Gly Asn Ala Ser Pro Ala Ser Gly
 275 280 285
 35 Ala Asn Pro Gly Ala Asn Gln Pro Gly Ser Ala Asp Asp Gln Ser Ser
 290 295 300
 40 Gly Gln Asn Asn Leu Gln Ser Gln Ile Met Asp Val Val Lys Glu Val
 305 310 315 320
 Val Gln Ile Leu Gln Gln Met Leu Ala Ala Gln Asn Gly Gly Ser Gln
 325 330 335
 45 Gln Ser Thr Ser Thr Gln Pro Met
 340

50 It is encoded by a DNA molecule having a nucleotide
 sequence corresponding SEQ. ID. No. 8 as follows:

ATGTCAGTCG GAAACATCCA GAGCCCGTCG AACCTCCCGG GTCTGCAGAA CCTGAACCTC 60
 55 AACACCAACA CCAACAGCCA GCAATCGGGC CAGTCCGTGC AAGACCTGAT CAAGCAGGTC 120
 GAGAAGGACA TCCTCAACAT CATCGCAGCC CTCGTGCAGA AGGCCGCACA GTCGGCGGGC 180
 60 GGCAACACCG GTAACACCGG CAACGCGCCG GCGAAGGACG GCAATGCCAA CGCGGGCGCC 240

- 25 -

```

AACGACCCGA GCAAGAACGA CCCGAGCAAG AGCCAGGCTC CGCAGTCGGC CAACAAGACC      300
GGCAACGTCG ACGACGCCAA CAACCAGGAT CCGATGCAAG CGCTGATGCA GCTGCTGGAA      360
5  GACCTGGTGA AGCTGCTGAA GGGCGCCCTG CACATGCAGC AGCCCGGCGG CAATGACAAG      420
GGCAACGGCG TGGGCGGTGC CAACGGCGCC AAGGGTGCCG GCGGCCAGGG CGGCCTGGCC      480
GAAGCGCTGC AGGAGATCGA GCAGATCCTC GCCCAGCTCG GCGGCGGCGG TGCTGGCGCC      540
10 GCGGCGCGCG GTGGCGGTGT CGGCGGTGCT GGTGGCGCGG ATGGCGGCTC CGGTGCGGGT      600
GGCGCAGGCG GTGCGAACGG CGCCGACGGC GGCAATGGCG TGAACGGCAA CCAGGCGAAC      660
15 GGCCCGCAGA ACGCAGGCGA TGTCAACGGT GCCAACGGCG CGGATGACGG CAGCGAAGAC      720
CAGGGCGGCC TCACCGGCGT GCTGCAAAAG CTGATGAAGA TCCTGAACGC GCTGGTGCAG      780
ATGATGCAGC AAGGCGGCCT CGGCGGCGGC AACCAGGCGC AGGGCGGCTC GAAGGGTGCC      840
20 GGCAACGCCT CGCCGGCTTC CGGCGCGAAC CCGGGCGCGA ACCAGCCCGG TTCGGCGGAT      900
GATCAATCGT CCGGCCAGAA CAATCTGCAA TCCAGATCA TGGATGTGGT GAAGGAGGTC      960
25 GTCCAGATCC TGCAGCAGAT GCTGGCGGCG CAGAACGGCG GCAGCCAGCA GTCCACCTCG     1020
ACGCAGCCGA TGTAAT                                         1035

```

30 Further information regarding the hypersensitive response elicitor polypeptide or protein derived from *Pseudomonas solanacearum* is set forth in Arlat, M., F. Van Gijsegem, J. C. Huet, J. C. Pemollet, and C. A. Boucher, "PopA1, a

35 Protein which Induces a Hypersensitive-like Response in Specific Petunia Genotypes, is Secreted via the Hrp Pathway of *Pseudomonas solanacearum*," EMBO J. 13:543-533 (1994), which is hereby incorporated by reference.

The hypersensitive response elicitor

40 polypeptide or protein from *Xanthomonas campestris* pv. *glycines* has an amino acid sequence corresponding to SEQ. ID. No. 9 as follows:

```

Thr Leu Ile Glu Leu Met Ile Val Val Ala Ile Ile Ala Ile Leu Ala
45 1           5           10           15

Ala Ile Ala Leu Pro Ala Tyr Gln Asp Tyr
                20           25

```

50

- 26 -

This sequence is an amino terminal sequence having 26 residues only from the hypersensitive response elicitor polypeptide or protein of *Xanthomonas campestris* pv. *glycines*. It matches with fimbrial subunit proteins determined in other *Xanthomonas campestris* pathovars.

The hypersensitive response elicitor polypeptide or protein from *Xanthomonas campestris pelargonii* is heat stable, protease sensitive, and has a molecular weight of 20kDa. It includes an amino acid sequence corresponding to SEQ. ID. No. 10 as follows:

	Ser	Ser	Gln	Gln	Ser	Pro	Ser	Ala	Gly	Ser	Glu	Gln	Gln	Leu	Asp	Gln
	1				5					10					15	
15	Leu	Leu	Ala	Met												
				20												

Isolation of *Erwinia carotovora* hypersensitive response elicitor protein or polypeptide is described in Cai, et al., "The RsmA Mutants of *Erwinia carotovora* subsp. *carotova* Strain Ecc71 Overexpress *hrpN_{Ecc}* and Elicit a Hypersensitive Reaction-Like Response in Tobacco Leaves," MPMI, 9(7):565-73 (1996), which is hereby incorporated by reference. The hypersensitive response elicitor protein or polypeptide for *Erwinia stewartii* is disclosed in Ahmad, et al., "Harpin is Not Necessary for the Pathogenicity of *Erwinia stewartii* on Maize," 8th Int'l. Cong. Molec. Plant-Microbe Interact, July 14-19, 1996 and Ahmad, et al., "Harpin is Not Necessary for the Pathogenicity of *Erwinia stewartii* on Maize," Ann. Mtg. Am. Phytopath. Soc., July 27-31, 1996, which are hereby incorporated by reference.

Hypersensitive response elicitor proteins or polypeptides from *Phytophthora parasitica*, *Phytophthora cryptogea*, *Phytophthora cinnamoni*, *Phytophthora capsici*, *Phytophthora megasperma*, and *Phytophthora citrophthora* are described in Kamoun, et al., "Extracellular Protein Elicitors from *Phytophthora*: Host-Specificity and

- 27 -

Induction of Resistance to Bacterial and Fungal
Phytopathogens," Molec. Plant-Microbe Interact., 6(1):15-
25 (1993), Ricci, et al., "Structure and Activity of
Proteins from Pathogenic Fungi *Phytophthora* Eliciting
5 Necrosis and Acquired Resistance in Tobacco," Eur. J.
Biochem., 183:555-63 (1989), Ricci, et al., "Differential
Production of Parasiticein, an Elicitor of Necrosis and
Resistance in Tobacco by Isolates of *Phytophthora*
paraticica," Plant Path., 41:298-307 (1992), Baillieul,
10 et al., "A New Elicitor of the Hypersensitive Response in
Tobacco: A Fungal Glycoprotein Elicits Cell Death,
Expression of Defense Genes, Production of Salicylic
Acid, and Induction of Systemic Acquired Resistance,"
Plant J., 8(4):551-60 (1995), and Bonnet, et al.,
15 "Acquired Resistance Triggered by Elicitins in Tobacco
and Other Plants," Eur. J. Plant Path., 102:181-92
(1996), which are hereby incorporated by reference.

The above elicitors are exemplary. Other
elicitors can be identified by growing fungi or bacteria
20 that elicit a hypersensitive response under which genes
encoding an elicitor are expressed. Cell-free
preparations from culture supernatants can be tested for
elicitor activity (i.e. local necrosis) by using them to
infiltrate appropriate plant tissues.

25 It is also possible to use fragments of the
above hypersensitive response elicitor polypeptides or
proteins as well as fragments of full length elicitors
from other pathogens, in the method of the present
invention.

30 Suitable fragments can be produced by several
means. In the first, subclones of the gene encoding a
known elicitor protein are produced by conventional
molecular genetic manipulation by subcloning gene
fragments. The subclones then are expressed *in vitro* or
35 *in vivo* in bacterial cells to yield a smaller protein or

- 28 -

a peptide that can be tested for elicitor activity according to the procedure described below.

As an alternative, fragments of an elicitor protein can be produced by digestion of a full-length elicitor protein with proteolytic enzymes like chymotrypsin or *Staphylococcus* proteinase A, or trypsin. Different proteolytic enzymes are likely to cleave elicitor proteins at different sites based on the amino acid sequence of the elicitor protein. Some of the fragments that result from proteolysis may be active elicitors of resistance.

In another approach, based on knowledge of the primary structure of the protein, fragments of the elicitor protein gene may be synthesized by using the PCR technique together with specific sets of primers chosen to represent particular portions of the protein. These then would be cloned into an appropriate vector for increased expression of a truncated peptide or protein.

An example of a suitable fragment is the popA1 fragment of the hypersensitive response elicitor polypeptide or protein from *Pseudomonas solanacearum*. See Arlat, M., F. Van Gijsegem, J.C. Huet, J.C. Pemollet, and C.A. Boucher, "PopA1, a Protein Which Induces a Hypersensitive-like Response in Specific *Petunia* Genotypes is Secreted via the Hrp Pathway of *Pseudomonas solanacearum*," EMBO J. 13:543-53 (1994), which is hereby incorporated by reference. As to *Erwinia amylovora*, a suitable fragment can be, for example, either or both the polypeptide extending between and including amino acids 1 and 98 of SEQ. ID. NO. 3 and the polypeptide extending between and including amino acids 137 and 204 of SEQ. ID. No. 3.

Variants may be made by, for example, the deletion or addition of amino acids that have minimal influence on the properties, secondary structure and

- 29 -

hydropathic nature of the polypeptide. For example, a polypeptide may be conjugated to a signal (or leader) sequence at the N-terminal end of the protein which co-translationally or post-translationally directs transfer of the protein. The polypeptide may also be conjugated to a linker or other sequence for ease of synthesis, purification or identification of the polypeptide.

The protein or polypeptide of the present invention is preferably produced in purified form (preferably at least about 60%, more preferably 80%, pure) by conventional techniques. Typically, the protein or polypeptide of the present invention is produced but not secreted into the growth medium of recombinant *E. coli*. Alternatively, the protein or polypeptide of the present invention is secreted into the growth medium. In the case of unsecreted protein, to isolate the protein, the *E. coli* host cell carrying a recombinant plasmid is propagated, homogenized, and the homogenate is centrifuged to remove bacterial debris. The supernatant is then subjected to heat treatment and the hypersensitive response elicitor protein is separated by centrifugation. The supernatant fraction containing the polypeptide or protein of the present invention is subjected to gel filtration in an appropriately sized dextran or polyacrylamide column to separate the proteins. If necessary, the protein fraction may be further purified by ion exchange or HPLC.

Alternatively, the hypersensitive response elicitor protein can be prepared by chemical synthesis using conventional techniques.

The DNA molecule encoding the hypersensitive response elicitor polypeptide or protein can be incorporated in cells using conventional recombinant DNA technology. Generally, this involves inserting the DNA molecule into an expression system to which the DNA

- 30 -

molecule is heterologous (i.e. not normally present).

The heterologous DNA molecule is inserted into the expression system or vector in proper sense orientation and correct reading frame. The vector contains the
5 necessary elements for the transcription and translation of the inserted protein-coding sequences.

U.S. Patent No. 4,237,224 to Cohen and Boyer, which is hereby incorporated by reference, describes the production of expression systems in the form of
10 recombinant plasmids using restriction enzyme cleavage and ligation with DNA ligase. These recombinant plasmids are then introduced by means of transformation and replicated in unicellular cultures including procaryotic organisms and eucaryotic cells grown in tissue culture.

15 Recombinant genes may also be introduced into viruses, such as vaccina virus. Recombinant viruses can be generated by transection of plasmids into cells infected with virus.

Suitable vectors include, but are not limited
20 to, the following viral vectors such as lambda vector system gt11, gt WES.tB, Charon 4, and plasmid vectors such as pBR322, pBR325, pACYC177, pACYC1084, pUC8, pUC9, pUC18, pUC19, pLG339, pR290, pKC37, pKC101, SV 40, pBluescript II SK +/- or KS +/- (see "Stratagene Cloning
25 Systems" Catalog (1993) from Stratagene, La Jolla, Calif, which is hereby incorporated by reference), pQE, pIH821, pGEX, pET series (see F.W. Studier et. al., "Use of T7 RNA Polymerase to Direct Expression of Cloned Genes,"
Gene Expression Technology vol. 185 (1990), which is
30 hereby incorporated by reference), and any derivatives thereof. Recombinant molecules can be introduced into cells via transformation, particularly transduction, conjugation, mobilization, or electroporation. The DNA sequences are cloned into the vector using standard
35 cloning procedures in the art, as described by Sambrook

- 31 -

et al., Molecular Cloning: A Laboratory Manual, Cold Springs Laboratory, Cold Springs Harbor, New York (1989), which is hereby incorporated by reference.

A variety of host-vector systems may be
5 utilized to express the protein-encoding sequence(s). Primarily, the vector system must be compatible with the host cell used. Host-vector systems include but are not limited to the following: bacteria transformed with bacteriophage DNA, plasmid DNA, or cosmid DNA;
10 microorganisms such as yeast containing yeast vectors; mammalian cell systems infected with virus (e.g., vaccinia virus, adenovirus, etc.); insect cell systems infected with virus (e.g., baculovirus); and plant cells infected by bacteria. The expression elements of these
15 vectors vary in their strength and specificities. Depending upon the host-vector system utilized, any one of a number of suitable transcription and translation elements can be used.

Different genetic signals and processing events
20 control many levels of gene expression (e.g., DNA transcription and messenger RNA (mRNA) translation).

Transcription of DNA is dependent upon the presence of a promotor which is a DNA sequence that directs the binding of RNA polymerase and thereby
25 promotes mRNA synthesis. The DNA sequences of eucaryotic promoters differ from those of procaryotic promoters. Furthermore, eucaryotic promoters and accompanying genetic signals may not be recognized in or may not function in a procaryotic system, and, further,
30 procaryotic promoters are not recognized and do not function in eucaryotic cells.

Similarly, translation of mRNA in procaryotes depends upon the presence of the proper procaryotic signals which differ from those of eucaryotes. Efficient
35 translation of mRNA in procaryotes requires a ribosome

- 32 -

binding site called the Shine-Dalgarno ("SD") sequence on the mRNA. This sequence is a short nucleotide sequence of mRNA that is located before the start codon, usually AUG, which encodes the amino-terminal methionine of the protein. The SD sequences are complementary to the 3'-end of the 16S rRNA (ribosomal RNA) and probably promote binding of mRNA to ribosomes by duplexing with the rRNA to allow correct positioning of the ribosome. For a review on maximizing gene expression, see Roberts and Lauer, Methods in Enzymology, 68:473 (1979), which is hereby incorporated by reference.

Promoters vary in their "strength" (i.e. their ability to promote transcription). For the purposes of expressing a cloned gene, it is desirable to use strong promoters in order to obtain a high level of transcription and, hence, expression of the gene. Depending upon the host cell system utilized, any one of a number of suitable promoters may be used. For instance, when cloning in *E. coli*, its bacteriophages, or plasmids, promoters such as the T7 phage promoter, *lac* promoter, *trp* promoter, *recA* promoter, ribosomal RNA promoter, the P_R and P_L promoters of coliphage lambda and others, including but not limited, to *lacUV5*, *ompF*, *bla*, *lpp*, and the like, may be used to direct high levels of transcription of adjacent DNA segments. Additionally, a hybrid *trp-lacUV5* (*tac*) promoter or other *E. coli* promoters produced by recombinant DNA or other synthetic DNA techniques may be used to provide for transcription of the inserted gene.

Bacterial host cell strains and expression vectors may be chosen which inhibit the action of the promoter unless specifically induced. In certain operations, the addition of specific inducers is necessary for efficient transcription of the inserted DNA. For example, the *lac* operon is induced by the

- 33 -

addition of lactose or IPTG (isopropylthio-beta-D-galactoside). A variety of other operons, such as *trp*, *pro*, etc., are under different controls.

Specific initiation signals are also required
5 for efficient gene transcription and translation in procaryotic cells. These transcription and translation initiation signals may vary in "strength" as measured by the quantity of gene specific messenger RNA and protein synthesized, respectively. The DNA expression vector,
10 which contains a promotor, may also contain any combination of various "strong" transcription and/or translation initiation signals. For instance, efficient translation in *E. coli* requires a Shine-Dalgarno (SD) sequence about 7-9 bases 5' to the initiation codon (ATG)
15 to provide a ribosome binding site. Thus, any SD-ATG combination that can be utilized by host cell ribosomes may be employed. Such combinations include but are not limited to the SD-ATG combination from the *cro* gene or the *N* gene of coliphage lambda, or from the *E. coli*
20 tryptophan E, D, C, B or A genes. Additionally, any SD-ATG combination produced by recombinant DNA or other techniques involving incorporation of synthetic nucleotides may be used.

Once the isolated DNA molecule encoding the
25 hypersensitive response elicitor polypeptide or protein has been cloned into an expression system, it is ready to be incorporated into a host cell. Such incorporation can be carried out by the various forms of transformation noted above, depending upon the vector/host cell system.
30 Suitable host cells include, but are not limited to, bacteria, virus, yeast, mammalian cells, insect, plant, and the like.

The method of the present invention can be utilized to treat seeds for a wide variety of plants to
35 impart pathogen resistance to the plants. Suitable seeds

- 34 -

are for plants which are dicots and monocots. More particularly, useful crop plants can include: rice, wheat, barley, rye, oats, cotton, sunflower, canola, peanut, corn, potato, sweet potato, bean, pea, chicory, lettuce, endive, cabbage, cauliflower, broccoli, turnip, radish, spinach, onion, garlic, eggplant, pepper, celery, carrot, squash, pumpkin, zucchini, cucumber, apple, pear, melon, strawberry, grape, raspberry, pineapple, soybean, tobacco, tomato, sorghum, and sugarcane. Examples of suitable ornamental plants are: rose, *Saintpaulia*, petunia, *Pelargonium*, poinsettia, chrysanthemum, carnation, and zinnia.

The method of imparting pathogen resistance to plants in accordance with the present invention is useful in imparting resistance to a wide variety of pathogens including viruses, bacteria, and fungi.

Resistance, *inter alia*, to the following viruses can be achieved by the method of the present invention: Tobacco mosaic virus, cucumber mosaic virus, potato x virus, potato y virus, and tomato mosaic virus.

Resistance, *inter alia*, to the following bacteria can also be imparted to plants in accordance with the present invention: *Pseudomonas solanacearum*, *Pseudomonas syringae* pv. *tabaci*, and *Xanthomonas campestris* pv. *pelargonii*.

Plants can be made resistant, *inter alia*, to the following fungi by use of the method of the present invention: *Fusarium oxysporum* and *Phytophthora infestans*.

The embodiment of the present invention involving applying the hypersensitive response elicitor polypeptide or protein to all or part of the plant seeds being treated can be carried out through a variety of procedures. Suitable application methods include high or low pressure spraying, injection, coating, dusting, and

- 35 -

immersion. Other suitable application procedures can be envisioned by those skilled in the art. Once treated with the hypersensitive response elicitor of the present invention, the seeds can be planted and cultivated using conventional procedures to produce plants. After plants have been propagated from seeds treated in accordance with the present invention, the plants may be treated with one or more applications of the hypersensitive response elicitor protein or polypeptide to enhance hypersensitive response induced resistance in the plants. See U.S. Patent Application Serial No. 08/475,775, which is hereby incorporated by reference. Such propagated plants, which are resistant to disease, may, in turn, be useful in producing seeds or propagules (e.g. cuttings) that produce resistant plants.

The hypersensitive response elicitor polypeptide or protein can be applied to plant seeds in accordance with the present invention alone or in a mixture with other materials.

A composition suitable for treating plant seeds in accordance with the present invention contains a hypersensitive response elicitor polypeptide or protein in a carrier. Suitable carriers include water, aqueous solutions, slurries, or dry powders. In this embodiment, the composition contains greater than 0.5 nM hypersensitive response elicitor polypeptide or protein.

Although not required, this composition may contain additional additives including fertilizer, insecticide, fungicide, nematocide, herbicide, and mixtures thereof. Suitable fertilizers include $(\text{NH}_4)_2\text{NO}_3$. An example of a suitable insecticide is Malathion. Useful fungicides include Captan.

Other suitable additives include buffering agents, wetting agents, coating agents, and abrading agents. These materials can be used to facilitate the

- 36 -

process of the present invention. In addition, the hypersensitive response elicitor polypeptide or protein can be applied to plant seeds with other conventional seed formulation and treatment materials, including clays and polysaccharides.

In the alternative embodiment of the present invention involving the use of transgenic seeds, a hypersensitive response elicitor polypeptide or protein need not be applied topically to the seeds. Instead, transgenic plants transformed with a DNA molecule encoding a hypersensitive response elicitor polypeptide or protein are produced according to procedures well known in the art, such as biolistics or *Agrobacterium* mediated transformation. Examples of suitable hypersensitive response elicitor polypeptides or proteins and the nucleic acid sequences for their encoding DNA are disclosed *supra*. As is conventional in the art, such transgenic plants would contain suitable vectors with various promoters including pathogen-induced promoters, and other components needed for transformation, transcription, and, possibly, translation. Such transgenic plants themselves could be grown under conditions effective to be imparted with pathogen resistance. In any event, once transgenic plants of this type are produced, transgenic seeds are recovered. These seeds can then be planted in the soil and cultivated using conventional procedures to produce plants. The plants are propagated from the planted transgenic seeds under conditions effective to impart pathogen resistance to the plants.

When transgenic plant seeds are used in accordance with the present invention, they additionally can be treated with the same materials (noted above) as are used to treat the seeds to which a hypersensitive response elicitor polypeptide or protein is applied.

- 37 -

These other materials, including hypersensitive response elicitors, can be applied to the transgenic plant seeds by high or low pressure spraying, injection, coating, dusting, and immersion. Similarly, transgenic plants additionally may be treated with one or more applications of the hypersensitive response elicitor to enhance hypersensitive response induced resistance in the plants. Such plants may also be treated with conventional plant treatment agents (e.g., insecticides, fertilizers, etc.). The transgenic plants of the present invention are useful in producing seeds or propagules (e.g. cuttings) from which disease resistant plants grow.

EXAMPLES

15

Example 1 - Effect of Treating Seeds with Hypersensitive Response Elicitor Protein

20

Marglobe tomato seeds were submerged in hypersensitive response elicitor protein (ca. 26 μ gm/ml) from *Erwinia amylovora* solution or buffer in beakers on day 0 for 24 hours at 28°C in a growth chamber. After soaking seeds in hypersensitive response elicitor protein from *Erwinia amylovora* or buffer, they were sown in germination pots with artificial soil on day 1. Seedlings were transplanted to individual pots at the two-true-leaf stage on day 12. After transplanting, some plants that arose from treated seed also were sprayed with hypersensitive response elicitor protein (ca. 13 μ gm/ml) from *Erwinia amylovora* (Treatments 3 and 4).

30

Tomato treated as noted in the preceding paragraph were inoculated with *Burkholderia* (*Pseudomonas*) *solanacearum* K60 strain (See Kelman, "The Relationship of Pathogenicity in *Pseudomonas solanacearum* to Colony Appearance on a Tetrazolium Medium," Phytopathology 44:693-95 (1954)) on day 23 by making vertical cuts

35

- 38 -

through the roots and potting medium of tomato plants (on a tangent 2 cm from the stem and 2 times/pot) and putting 10 ml (5×10^8 cfu/ml) suspension into the soil.

5 The above procedure involved use of 10 seeds treated with hypersensitive response elicitor protein from *Erwinia amylovora* per treatment.

Treatments:

- 10 1. Seeds soaked in hypersensitive response elicitor protein from *Erwinia amylovora* (ca. 26 μ g/ml).
2. Seeds soaked in buffer (5mM KPO₄, pH 6.8).
- 15 3. Seeds soaked in hypersensitive response elicitor protein from *Erwinia amylovora* (ca. 26 μ g/ml) and seedlings sprayed with hypersensitive response elicitor protein from *Erwinia amylovora* (ca. 13 μ g/ml) at transplanting.
- 20 4. Seeds soaked in buffer and seedlings sprayed with hypersensitive response elicitor protein from *Erwinia amylovora* (ca. 13 μ g/ml) at transplanting.

25 The results of these treatments are set forth in Tables 1-4.

**Table 1 - Infection Data - 28 Days After Seed
Treatment and 5 Days After Inoculation**

5			Number of Plants of Given Disease Rating*					
	Treatm.	Plants	0	1	2	3	4	5
	1	10	10	0	0	0	0	0
	2	10	9	1	0	0	0	0
	3	10	9	1	0	0	0	0
10	4	10	10	0	0	0	0	0

*** Disease Scale:**

- Grade 0: No symptoms
 Grade 1: One leaf partially wilted.
 Grade 2: 2-3 leaves wilted.
 Grade 3: All except the top 2-3 leaves wilted.
 Grade 4: All leaves wilted.
 Grade 5: Plant Dead

**Table 2 - Infection Data - 31 Days After Seed
Treatment and 8 Days After Inoculation**

25			Number of Plants of Given Disease Rating*					
	Treatm.	Plants	0	1	2	3	4	5
	1	10	6	4	0	0	0	0
	2	10	4	3	2	1	0	0
	3	10	8	2	0	0	0	0
30	4	10	7	2	1	0	0	0

- 40 -

Table 3 - Infection Data - 35 Days After Seed Treatment and 12 Days After Inoculation

		Number of Plants of Given Disease Rating*					
Treatm.	Plants	0	1	2	3	4	5
1	10	5	3	0	1	1	0
2	10	1	3	3	2	1	0
3	10	4	3	3	0	0	0
4	10	3	3	3	1	0	0

Table 4 - Disease Indices of Seed Treatment With Hypersensitive Response Elicitor Protein

Treatment		Inoculation	Disease Index (%)*		
Day 0	Day 14	Day 23	Day 28	Day 31	Day 35
1. Hypersensitive response elicitor protein seed soak		Inoculate	0	8	20
2. Buffer seed soak		Inoculate	2	20	38
3. Hypersensitive response elicitor protein seed soak	Spray Hypersensitive response elicitor protein	Inoculate	2	4	18
4. Buffer seed soak	Spray Hypersensitive response elicitor protein	Inoculate	0	8	24

* The Disease Index was determined using the procedure set forth in N.N. Winstead, et al., "Inoculation Techniques for Evaluating Resistance to *Pseudomonas Solanacearum*," *Phytopathology* 42:628-34 (1952), particularly at page 629.

- 41 -

The above data shows that the hypersensitive response elicitor protein was more effective than buffer as a seed treatment in reducing disease index and was as effective as spraying leaves of young plants with
 5 *hypersensitive response elicitor protein.*

Example 2 - Effect of Treating Tomato Seeds With
 Hypersensitive Response Elicitor Protein
 From pCPP2139 Versus pCPP50 Vector On
 10 Southern Bacteria Wilt Of Tomato

Marglobe tomato seeds were submerged in hypersensitive response elicitor protein from pCPP2139 or in pCPP50 vector solution (1:50, 1:100 and 1:200) in
 15 beakers on day 0 for 24 hours at 28°C in a growth chamber. After soaking seeds in hypersensitive response elicitor protein or vector, they were sown in germination pots with artificial soil on day 0. Ten uniform
 20 appearing plants were chosen randomly from each of the following treatments:

	Treatment Content	Strain	Dilution	Harpin
25	1.	DH5 α (pCPP2139)	1:50	8 μ g/ml
	2.	DH5 α (pCCP50)	1:50	0
	3.	DH5 α (pCPP2139)	1:100	4 μ g/ml
	4.	DH5 α (pCPP50)	1:100	0
	5.	DH5 α (pCPP2139)	1:200	2 μ g/ml
30	6.	DH5 α (pCPP50)	1:200	0

The resulting seedlings were inoculated with *Pseudomonas solanacearum* K60 by dipping the roots of tomato seedling
 35 plants for about 30 seconds in a 40 ml (1×10^8 cfu/ml) suspension. The seedlings were then transplanted into the pots with artificial soil on day 12.

The results of these treatments are set forth
 40 in Tables 5-8.

**Table 5 - 16 Days After Seed Treatment and
3 Days After Inoculation**

5			Number of Plants of Given Disease Rating*					
	Treatm.	Plants	0	1	2	3	4	5
10	1	10	7	3	0	0	0	0
	2	10	5	5	0	0	0	0
	3	10	6	4	0	0	0	0
	4	10	6	4	0	0	0	0
	5	10	7	4	0	0	0	0
	6	10	4	6	0	0	0	0

**Table 6 - 19 Days After Seed Treatment and
6 Days After Inoculation**

15			Number of Plants of Given Disease Rating*					
	Treatm.	Plants	0	1	2	3	4	5
20	1	10	6	0	0	0	0	0
	2	10	2	0	2	2	1	3
	3	10	2	0	2	0	2	4
	4	10	3	1	2	0	2	2
	5	10	2	1	0	2	2	3
25	6	10	1	0	1	1	3	4

- 43 -

Table 7 - 21 Days After Seed Treatment and
8 Days After Inoculation

		Number of Plants of Given Disease Rating*							
5	Treatm.	Plants	0	1	2	3	4	5	
	1	10	6	0	0	0	1	3	
	2	10	2	0	0	1	3	4	
	3	10	2	0	0	2	2	3	
	4	10	3	0	0	2	2	3	
	10	5	10	2	0	0	0	4	4
		6	10	1	0	1	2	1	5

Table 8 - Disease Indices of Seed Treatment
With Hypersensitive Response Elicitor and Vector

	Treatment		Disease Index (%)		
	Day 0	Day 12	Day 15	Day 18	Day 20
20	Hypersensitive response elicitor protein seed dip (1:50)	inoculate	6.0	32.0	38.0
	Vector seed dip (1:50)	inoculate	10.0	58.0	70.0
25	Hypersensitive response elicitor protein seed dip (1:100)	inoculate	8.0	64.0	68.0
30	Vector seed dip (1:100)	inoculate	8.0	46.0	58.0
	Hypersensitive response elicitor protein seed dip (1:200)	inoculate	6.0	60.00	72.0
35	Vector seed dip (1:200)	inoculate	12.0	74.0	74.0

The above data shows that the hypersensitive response elicitor protein is much more effective than the vector solution in preventing Tomato Southern Bacteria Wilt.

- 44 -

**Example 3 - Effect of Treating Tomato Seeds With
Hypersensitive Response Elicitor Protein
From pCPP2139 Versus pCPP50 Vector On
Tomato Southern Bacteria Wilt**

5 Marglobe tomato seeds were submerged in
hypersensitive response elicitor protein from pCPP2139 or
in pCPP50 vector solution (1:50, 1:100 and 1:200) in
beakers on day 0 for 24 hours at 28°C in a growth
10 chamber. After soaking seeds in the hypersensitive
response elicitor protein or vector, the seeds were sown
in germination pots with artificial soil on day 1. Ten
uniform appearing plants were chosen randomly from each
of the following treatments:

15	Treatment	Strain	Dilution	Hypersensitive Response Elicitor Content
20	1.	DH5 α (pCPP2139)	1:50	8 μ g/ml
	2.	DH5 α (pCPP50)	1:50	0
	3.	DH5 α (pCPP2139)	1:100	4 μ g/ml
	4.	DH5 α (pCPP50)	1:100	0
	5.	DH5 α (pCPP2139)	1:200	2 μ g/ml
25	6.	DH5 α (pCPP50)	1:200	0

The resulting seedlings were inoculated with *Pseudomonas*
solanacearum K60 by dipping the roots of tomato seedling plants
30 for about 30 seconds in a 40 ml (1 X 10⁶ cfu/ml) suspension.
The seedlings were then transplanted into the pots with
artificial soil on day 12.

The results of these treatments are set forth in
Tables 9-12.

35

- 45 -

Table 9 - 16 Days After Seed Treatment and
3 Days After Inoculation

		Number of Plants of Given Disease Rating*						
5	Treatm.	Plants	0	1	2	3	4	5
	1	10	8	2	0	0	0	0
	2	10	7	3	0	0	0	0
	3	10	7	3	0	0	0	0
	4	10	7	3	0	0	0	0
	5	10	8	2	0	0	0	0
10	6	10	7	3	0	0	0	0

Table 10 - 19 Days After Seed Treatment and
6 Days After Inoculation

		Number of Plants of Given Disease Rating*						
20	Treatm.	Plants	0	1	2	3	4	5
	1	10	5	0	0	1	2	2
	2	10	1	0	1	2	3	3
	3	10	4	1	0	0	2	3
	4	10	2	0	2	1	2	3
	5	10	1	0	1	1	4	3
25	6	10	1	0	0	2	4	3

- 46 -

**Table 11 - 21 Days After Hypersensitive Response
Elicitor Protein Seed Treatment and
8 Days After Inoculation**

Treatm.	Plants	Number of Plants of Given Disease Rating*					
		0	1	2	3	4	5
1	10	5	0	0	0	2	3
2	10	2	0	2	0	2	4
3	10	5	0	0	0	2	3
4	10	2	0	2	0	2	4
5	10	1	0	1	0	2	6
6	10	1	0	0	0	2	7

**Table 12 - Disease Indices of Seed Treatment
With Hypersensitive Response Elicitor Protein and Vector**

Day 1	Day 13	Day 16	Day 19	Day 21
Hypersensitive response elicitor protein seed dip (1:50)	inoculate	4.0	42.0	46.0
Vector seed dip (1:50)	inoculate	6.0	70.0	64.0
Hypersensitive response elicitor protein seed dip (1:100)	inoculate	6.0	48.0	46.0
Vector seed dip (1:100)	inoculate	6.0	60.0	64.0
Hypersensitive response elicitor protein seed dip (1:200)	inoculate	4.0	72.0	80.0
Vector seed dip (1:200)	inoculate	6.0	74.0	86.0

The above data shows that the hypersensitive response elicitor protein is much more effective in preventing Tomato Southern Bacteria Wilt.

- 47 -

**Example 4 - Effect of Treating Tomato Seeds With
Hypersensitive Response Elicitor Protein
From pCPP2139 Versus pCPP50 Vector On
Southern Bacteria Wilt Of Tomato**

5 Marglobe tomato seeds were submerged in
hypersensitive response elicitor protein from pCPP2139 or
in pCPP50 vector solution (1:25, 1:50 and 1:100) in
beakers on day 0 for 24 hours at 28°C in a growth
10 chamber. After soaking seeds in hypersensitive response
elicitor protein or vector, they were sown in germination
pots with artificial soil on day 1. Ten uniform
appearing plants were chosen randomly from each of the
following treatments:

	Treatment Content	Strain	Dilution	Harpin
	1.	DH5α (pCPP2139)	1:25	16 µg/ml
20	2.	DH5α (pCPP50)	1:25	0
	3.	DH5α (pCPP2139)	1:50	8 µg/ml
	4.	DH5α (pCPP50)	1:50	0
	5.	DH5α (pCPP2139)	1:100	2 µg/ml
25	6.	DH5α (pCPP50)	1:100	0

The resulting seedlings were inoculated with *Pseudomonas*
solanacearum K60 by dipping the roots of tomato seedling
plants for about 30 seconds in a 40 ml (1×10^8 cfu/ml)
30 suspension. The seedlings were then transplanted into
the pots with artificial soil on day 14.

The results of these treatments are set forth
in Tables 13-16.

35

- 48 -

**Table 13 - 19 Days After Seed Treatment and
4 Days After Inoculation**

		Number of Plants of Given Disease Rating*						
5	Treatm.	Plants	0	1	2	3	4	5
	1	10	8	2	0	0	0	0
	2	10	5	2	2	1	0	0
	3	10	9	1	0	0	0	0
	4	10	5	2	1	2	0	0
	10	5	10	5	3	1	1	0
6		10	6	1	2	1	0	0

**Table 14 - 21 Days After Seed Treatments and
6 Days After Inoculation**

		Number of Plants of Given Disease Rating*					
Treatm.	Plants	0	1	2	3	4	5
1	10	6	3	0	0	1	0
2	10	3	2	1	0	0	0
3	10	6	3	1	0	0	0
4	10	3	2	1	2	2	0
5	10	5	1	2	2	0	0
6	10	3	1	3	2	1	0

- 49 -

**Table 15 - 23 Days After Seed Treatment and
8 Days After Inoculation**

		Number of Plants of Given Disease Rating*							
5	Treatm.	Plants	0	1	2	3	4	5	
	1	10	7	2	0	0	0	1	
	2	10	2	2	2	3	0	1	
	3	10	7	2	0	1	0	0	
	4	10	2	1	2	3	0	2	
	10	5	10	3	1	2	3	0	1
		6	10	2	2	2	3	0	1

**Table 16 - Disease Indices of Seed Treatment
With Hypersensitive Elicitor Protein and Vector**

Treatment		Disease Index (%)		
Day 1	Day 15	Day 19	Day 21	Day 23
20 Hypersensitive response elicitor protein seed dip (1:25)	inoculate	4.0	14.0	14.0
25 Vector seed dip (1:25)	inoculate	18.0	28.0	40.0
Hypersensitive response elicitor protein seed dip (1:50)	inoculate	2.0	10.0	10.0
30 Vector seed dip (1:50)	inoculate	20.0	36.0	48.0
Hypersensitive response elicitor protein seed dip (1:100)	inoculate	16.0	22.0	38.0
35 Vector seed dip (1:100)	inoculate	16.0	34.0	40.0

40 The above data shows that the hypersensitive response elicitor protein is much more effective than the vector solution in preventing Tomato Southern Bacteria

- 50 -

Wilt. A hypersensitive response protein concentration of 1:50 is particularly effective in disease control.

5 **Example 5 - Effect of Treating Tomato Seeds With
Hypersensitive Response Elicitor Protein
From pCPP2139 Versus pCPP50 Vector On
Southern Bacteria Wilt Of Tomato**

10 Marglobe tomato seeds were submerged in
hypersensitive response elicitor protein from pCPP2139 or
pCPP50 vector solution (1:25, 1:50 and 1:100) in beakers
on day 0 for 24 hours at 28°C in a growth chamber. After
soaking seeds in hypersensitive response elicitor protein
or vector, they were sown in germination pots with
15 artificial soil on day 1. Ten uniform appearing plants
were chosen randomly from each of the following
treatments:

20	Treatment Content	Strain	Dilution	Harpin
	1.	DH5α (pCPP2139)	1:25	16 µg/ml
	2.	DH5α (pCPP50)	1:25	0
	3.	DH5α (pCPP2139)	1:50	8 µg/ml
25	4.	DH5α (pCPP50)	1:50	0
	5.	DH5α (pCPP2139)	1:100	4 µg/ml
	6.	DH5α (pCPP50)	1:100	0

30 The resulting seedlings were inoculated with *Pseudomonas solanacearum* K60 by dipping the roots of tomato seedling plants for about 30 seconds in a 40 ml (1×10^6 cfu/ml) suspension. The seedlings were then transplanted into the pots with artificial soil on day 14.

35 The results of these treatments are set forth in Tables 17-20.

- 51 -

**Table 17 - 19 Days After Seed Treatment and
4 Days After Inoculation**

		Number of Plants of Given Disease Rating*						
5	Treatm.	Plants	0	1	2	3	4	5
	1	10	8	2	0	0	0	0
	2	10	6	3	1	0	0	0
	3	10	9	1	0	0	0	0
	4	10	6	4	0	0	0	0
10	5	10	6	2	1	1	0	0
	6	10	6	4	0	0	0	0

**Table 18 - 21 Days After Seed Treatment and
6 Days After Inoculation**

		Number of Plants of Given Disease Rating*						
20	Treatm.	Plants	0	1	2	3	4	5
	1	10	7	1	1	1	0	0
	2	10	3	3	2	2	0	0
	3	10	8	2	0	0	0	0
	4	10	3	3	2	2	0	0
	5	10	6	1	1	2	0	0
	6	10	3	2	3	1	1	0

- 52 -

**Table 19 - 23 Days After Seed Treatment and
8 Days After Inoculation**

		Number of Plants of Given Disease Rating*							
5	Treatm.	Plants	0	1	2	3	4	5	
	1	10	7	0	2	1	0	0	
	2	10	3	1	2	3	0	1	
	3	10	8	1	0	1	0	0	
	4	10	3	3	1	2	0	1	
	10	5	10	3	3	0	2	1	1
		6	10	3	2	0	3	0	2

**Table 20 - Disease Indices of Seed Treatment
With Hypersensitive Response Elicitor Protein and Vector**

	Treatment		Disease Index (%)		
	Day 0	Day 15	Day 19	Day 21	Day 23
20	Hypersensitive response elicitor protein seed dip (1:25)	inoculate	4.0	12.0	14.0
25	Vector seed dip (1:25)	inoculate	10.0	26.0	38.0
30	Hypersensitive response elicitor protein seed dip (1:50)	inoculate	2.0	4.0	8.0
	Vector seed dip (1:50)	inoculate	8.0	26.0	32.0
35	Hypersensitive response elicitor protein seed dip (1:100)	inoculate	14.0	18.0	36.0
	Vector seed dip (1:100)	inoculate	8.0	30.0	42.0

The above data shows that the hypersensitive response elicitor protein is much more effective than the

- 53 -

vector solution in preventing Tomato Southern Bacteria Wilt. A hypersensitive response elicitor protein concentration of 1:50 is more effective in disease control.

5

Example 6 - Treating Rice Seeds with Hypersensitive Response Elicitor Protein to Reduce Rice Stem Rot

10 Rice seeds (variety, M-202) were submerged in two gallons of hypersensitive response elicitor protein solution at a concentration of 20 μ g for 24 hours at room temperature. Rice seeds submerged in the same solution without hypersensitive response elicitor protein were
15 used as a control. After soaking, the seeds were sown in a rice field by air plane spray. There were four replicates for both hypersensitive response elicitor protein and control treatment. The lot size of each replicate is 150 Ft². The design of each plot was
20 completely randomized, and each plot had substantial level contamination of *Sclerotium oryzae*. Three months after sowing, stem rot was evaluated according to the following rating scale: Scale 1 = no disease, 2 = disease present on the exterior of the leaf sheath, 3 =
25 disease penetrates leaf sheath completely but is not present on culm, 4 = disease is present on culm exterior but does not penetrate to interior of culm, and 5 = disease penetrates to interior of culm. 40 plants from each replicate were sampled and assessed for the disease
30 incidence and severity. From Table 21, it is apparent that treating seeds with hypersensitive response elicitor reduced both disease incidence and severity. More particularly, regarding incidence, 67% of the plants were infected by stem rot for the control treatment, however,
35 only 40% plants were infected for the hypersensitive response elicitor protein treatment. As to severity, the disease index* for the hypersensitive response elicitor

- 54 -

protein treatment was 34% and 60% for the control. Accordingly, treating rice seed with hypersensitive response elicitor protein resulted in a significant reduction of stem rot disease. The hypersensitive response elicitor protein-induced resistance in rice can last a season long. In addition to disease resistance, it was also observed that hypersensitive response elicitor protein-treated rice had little or no damage by army worm (*Spodoptera praefica*). In addition, the treated plants were larger and had deeper green color than the control plants.

Table 21 - Incidence and Severity of Stem Rot (*Schlerotium oryzae*) on Rice, M-202

Treatment	% plants given disease rating					Disease index(%) (severity)
	1	2	3	4	5	
Harpin 20 µg/ml	60	5	8	18	10	34
Control	33	5	18	28	18	60

*Disease Index (%) for the harpin treatment

$$= \frac{1 \times 60 + 2 \times 5 + 3 \times 8 + 4 \times 18 + 5 \times 10}{5 \times 100} \times 100/100$$

*Disease Index (%) for the control treatment

$$= \frac{1 \times 33 + 2 \times 5 + 3 \times 18 + 4 \times 28 + 5 \times 18}{5 \times 100 \times 100/100} \times 100/100$$

- 55 -

Example 7 - Effect of Treating Onion Seed with
Hypersensitive Response Elicitor Protein
on the Development of Onion Smut Disease
(*Urocystis cepulae*) and On Seedling
Emergence

Onion seed, variety Pennant, (Seed Lot# 64387),
obtained from the Crookham Co., Caldwell, ID 83606,
treated with hypersensitive response elicitor protein or
a control was planted in a natural organic or "muck"
soil. Some of the seedlings that grew from the sown seed
were healthy, some had lesions characteristic of the
Onion Smut disease, and some of the sown seed did not
produce seedlings that emerged from the soil. Thus, the
effect of treating onion seed with various concentrations
of hypersensitive response elicitor protein was
determined.

Naturally infested muck soil was obtained from
a field in Oswego County, NY, where onions had been grown
for several years and where the Onion Smut disease
commonly had been problematic. Buckets of muck (5-gallon
plastic) were stored at 4°C until used. The soil was
mixed, sieved, and put in plastic flats 10 inches wide,
20 inches long, and 2 inches deep for use in the tests
described. Based on preliminary experiments, the soil
contained many propagules of the Onion Smut fungus,
Urocystis cepulae, such that when onion seed was sown in
the soil, smut lesions developed on many of the seedlings
that emerged from the soil. In addition, the soil
harbored other microorganisms, including those that cause
the "damping-off" disease. Among the several fungi that
cause damping off are *Pythium*, *Fusarium*, and *Rhizoctonia*
species.

The hypersensitive response elicitor protein
encoded by the *hrpN* gene of *Erwinia amylovora* was used to
treat seeds. It was produced by fermentation of the
cloned gene in a high-expression vector in *E. coli*.
Analysis of the cell-free elicitor preparation by high-

- 56 -

pressure liquid chromatography indicated its hypersensitive response elicitor protein content and on that basis appropriate dilutions were prepared in water. Seeds were soaked in a beaker containing hypersensitive response elicitor protein concentrations of 0, 5, 25, and 50 $\mu\text{gm/ml}$ of hypersensitive response elicitor protein for 24 hours. They were removed, dried briefly on paper towels, and sown in the muck soil. Treated seed was arranged by row, 15 seeds in each row for each treatment; each flat contained two replicates, and there were six replicates. Thus, a total of 90 seeds were treated with each concentration of hypersensitive response elicitor protein. The flats containing the seeds were held in a controlled environment chamber operating at 60°F (15.6°C), with a 14-hour day /10-hour night. Observations were made on seedling emergence symptoms (smut lesions). The data were recorded 23 days after sowing.

The effect of soaking onion seed in different concentrations of hypersensitive response elicitor protein on emergence of onion seedlings and on the incidence of onion smut is shown in Table 22. Only slight differences in emergence were noted, suggesting that there is no significant effect of treating with hypersensitive response elicitor protein at the concentrations used. Among the seedlings that emerged, substantially more of the seeds that received no hypersensitive response elicitor protein exhibited symptoms of Onion Smut than seedlings that grew from seed that had been treated with hypersensitive response elicitor protein. Treating seed with 25 $\mu\text{gm/ml}$ of hypersensitive response elicitor protein was the most effective concentration tested in reducing Onion Smut. Thus, this example demonstrates that treating onion seed with hypersensitive response elicitor protein reduces the Onion Smut disease.

Table 22 - Effect of Treating Onion Seed With Hypersensitive Response Elicitor Protein (i.e. Harpin) on the Development of Onion Smut Disease (*Urocystis cepulae*).

Treatment harpin ($\mu\text{g/ml}$)	Mean Seedlings Emerged (of 15)	Mean Percent Emerged	<u>Emerged</u>	
			Percent Healthy	Percent with Smut
0	5.00	33.3	20.0	80.0
5	3.67	24.4	40.9	59.1
25	4.33 ¹	28.8	50.0	46.2
50	4.17	27.7	44.0	56.0

¹ One seedling emerged then died.

Example 8 - Effect of Treating Tomato Seed with Hypersensitive Response Elicitor Protein on the Development of Bacterial Speck of Tomato (*Pseudomonas syringae* pv. *tomato*)

Tomato seed, variety New Yorker (Seed lot# 2273-2B), obtained from Harris Seeds, Rochester, NY, were treated with four concentrations of hypersensitive response elicitor protein (including a no-elicitor protein, water-treated control) and planted in peatlite soil mix. After 12 days and when the seedlings were in the second true-leaf stage, they were inoculated with the Bacterial Speck pathogen. Ten days later, the treated and inoculated plants were evaluated for extent of infection. Thus, the effect of treating tomato seed with various concentrations of hypersensitive response elicitor protein on resistance to *Pseudomonas syringae* pv. *tomato* was determined.

The hypersensitive response elicitor protein encoded by the *hrpN* gene of *Erwinia amylovora* was used to treat seeds. It was produced by fermentation of the cloned gene in a high-expression vector in *E. coli*.

- 58 -

Analysis of the cell-free elicitor preparation by high-pressure liquid chromatography indicated its hypersensitive response elicitor protein content and, on that basis, appropriate dilutions were prepared in water.

5 Seeds were soaked in a beaker containing hypersensitive response elicitor protein concentrations of 0, 5, 10, and 20 $\mu\text{g}/\text{ml}$ of hypersensitive response elicitor protein for 24 hours. They were removed, dried briefly on paper towels, and sown. The soil was a mixture of peat and

10 Pearlite™ in plastic flats 10 inches wide, 20 inches long, and 2 inches deep. Treated seed was arranged by row, 6 seeds in each row for each treatment; each flat contained two replicates, and there were four replicates and thus a total of 24 seeds that were treated with each

15 concentration of hypersensitive response elicitor protein. The flats containing the seeds were held in a controlled environment chamber operating at 75°F (25°C), with a 14-hour day/10-hour night.

When twelve-days old, the tomato seedlings were

20 inoculated with 10^8 colony forming units/ml of the pathogen, applied as a foliar spray. The flats containing the seedlings were covered with a plastic dome for 48 hours after inoculation to maintain high humidity. Observations were made on symptom severity using a

25 rating scale of 0-5. The rating was based on the number of lesions that developed on the leaflets and the cotyledons and on the relative damage caused to the plant parts by necrosis that accompanied the lesions. The cotyledons and (true) leaflets were separately rated for

30 disease severity 11 days after inoculation

The effect of soaking tomato seed in different concentrations of hypersensitive response elicitor protein (i.e. harpin) on the development of Bacterial Speck on leaflets and cotyledons of tomato is shown in

35 Table 23. The seedlings that grew from seed treated with the highest amount of hypersensitive response elicitor

- 59 -

protein tested (20 $\mu\text{gm/ml}$) had fewer diseased leaflets and cotyledons than the treatments. The water-treated control seedlings did not differ substantially from the plants treated with the two lower concentrations of hypersensitive response elicitor protein. Considering the disease ratings, the results were similar. Only plants treated with the highest concentration of hypersensitive response elicitor protein had disease ratings that were less than those of the other treatments. This example demonstrates that treatment of tomato seed with hypersensitive response elicitor protein reduces the incidence and severity of Bacterial Speck of tomato.

Table 23 - Effect of Treating Tomato Seed With Hypersensitive Response Elicitor Protein (i.e. Harpin) on the Subsequent Development of Bacterial Speck Disease (*Pseudomonas syringae* pv. *tomato*) on Tomato Cotyledons and Tomato Leaflets

Treatment Harpin ($\mu\text{g/ml}$)	Cotyledons			Leaflets		
	Mean Diseased	Percent Diseased	Disease Rating	Mean Diseased	Percent Diseased	Disease Rating
0	6.0/9.0	66.6	0.8	25.8/68.8	37.5	0.5
5	5.3/7.3	72.4	0.8	22.5/68.0	37.5	0.5
10	5.8/8.0	72.3	0.8	25.5/66.0	38.6	0.5
20	5.3/8.5	61.8	0.6	23.8/73.5	32.3	0.4

Although the invention has been described in detail for the purpose of illustration, it is understood that such detail is solely for that purpose, and variations can be made therein by those skilled in the art without departing from the spirit and scope of the invention which is defined by the following claims.

- 60 -

SEQUENCE LISTING

(1) GENERAL INFORMATION:

- (i) APPLICANT: Cornell Research Foundation, Inc.
- (ii) TITLE OF INVENTION: HYPERSENSITIVE RESPONSE INDUCED
RESISTANCE IN PLANTS BY SEED
TREATMENT
- (iii) NUMBER OF SEQUENCES: 10
- (iv) CORRESPONDENCE ADDRESS:
 - (A) ADDRESSEE: Nixon, Hargrave, Devans & Doyle LLP
 - (B) STREET: P.O. Box 1051, Clinton Square
 - (C) CITY: Rochester
 - (D) STATE: New York
 - (E) COUNTRY: U.S.A.
 - (F) ZIP: 14603
- (v) COMPUTER READABLE FORM:
 - (A) MEDIUM TYPE: Floppy disk
 - (B) COMPUTER: IBM PC compatible
 - (C) OPERATING SYSTEM: PC-DOS/MS-DOS
 - (D) SOFTWARE: PatentIn Release #1.0, Version #1.30
- (vi) CURRENT APPLICATION DATA:
 - (A) APPLICATION NUMBER:
 - (B) FILING DATE:
 - (C) CLASSIFICATION:
- (vii) PRIOR APPLICATION DATA:
 - (A) APPLICATION NUMBER: US 60/033,230
 - (B) FILING DATE: 05-DEC-1996
- (viii) ATTORNEY/AGENT INFORMATION:
 - (A) NAME: Goldman, Michael L.
 - (B) REGISTRATION NUMBER: 30,727
 - (C) REFERENCE/DOCKET NUMBER: 19603/1202
- (ix) TELECOMMUNICATION INFORMATION:
 - (A) TELEPHONE: (716) 263-1304
 - (B) TELEFAX: (716) 263-1600

(2) INFORMATION FOR SEQ ID NO:1:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 338 amino acids
 - (B) TYPE: amino acid
 - (C) STRANDEDNESS:
 - (D) TOPOLOGY: linear

- 61 -

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

```

Met Gln Ile Thr Ile Lys Ala His Ile Gly Gly Asp Leu Gly Val Ser
1      5      10      15
Gly Leu Gly Ala Gln Gly Leu Lys Gly Leu Asn Ser Ala Ala Ser Ser
20      25      30
Leu Gly Ser Ser Val Asp Lys Leu Ser Ser Thr Ile Asp Lys Leu Thr
35      40      45
Ser Ala Leu Thr Ser Met Met Phe Gly Gly Ala Leu Ala Gln Gly Leu
50      55      60
Gly Ala Ser Ser Lys Gly Leu Gly Met Ser Asn Gln Leu Gly Gln Ser
65      70      75      80
Phe Gly Asn Gly Ala Gln Gly Ala Ser Asn Leu Leu Ser Val Pro Lys
85      90      95
Ser Gly Gly Asp Ala Leu Ser Lys Met Phe Asp Lys Ala Leu Asp Asp
100     105     110
Leu Leu Gly His Asp Thr Val Thr Lys Leu Thr Asn Gln Ser Asn Gln
115     120     125
Leu Ala Asn Ser Met Leu Asn Ala Ser Gln Met Thr Gln Gly Asn Met
130     135     140
Asn Ala Phe Gly Ser Gly Val Asn Asn Ala Leu Ser Ser Ile Leu Gly
145     150     155     160
Asn Gly Leu Gly Gln Ser Met Ser Gly Phe Ser Gln Pro Ser Leu Gly
165     170     175
Ala Gly Gly Leu Gln Gly Leu Ser Gly Ala Gly Ala Phe Asn Gln Leu
180     185     190
Gly Asn Ala Ile Gly Met Gly Val Gly Gln Asn Ala Ala Leu Ser Ala
195     200     205
Leu Ser Asn Val Ser Thr His Val Asp Gly Asn Asn Arg His Phe Val
210     215     220
Asp Lys Glu Asp Arg Gly Met Ala Lys Glu Ile Gly Gln Phe Met Asp
225     230     235     240
Gln Tyr Pro Glu Ile Phe Gly Lys Pro Glu Tyr Gln Lys Asp Gly Trp
245     250     255
Ser Ser Pro Lys Thr Asp Asp Lys Ser Trp Ala Lys Ala Leu Ser Lys
260     265     270
Pro Asp Asp Asp Gly Met Thr Gly Ala Ser Met Asp Lys Phe Arg Gln
275     280     285
Ala Met Gly Met Ile Lys Ser Ala Val Ala Gly Asp Thr Gly Asn Thr
290     295     300

```

- 62 -

Asn Leu Asn Leu Arg Gly Ala Gly Gly Ala Ser Leu Gly Ile Asp Ala
305 310 315 320

Ala Val Val Gly Asp Lys Ile Ala Asn Met Ser Leu Gly Lys Leu Ala
325 330 335

Asn Ala

(2) INFORMATION FOR SEQ ID NO:2:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 2141 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

CGATTTTACC	CGGGTGAACG	TGCTATGACC	GACAGCATCA	CGGTATTCGA	CACCGTTACG	60
GCGTTTATGG	CCGCGATGAA	CCGGCATCAG	GCGGCGCGCT	GGTCGCCGCA	ATCCGGCGTC	120
GATCTGGTAT	TTCAGTTTGG	GGACACCGGG	CGTGAACCTCA	TGATGCAGAT	TCAGCCGGGG	180
CAGCAATATC	CCGGCATGTT	GCGCACGCTG	CTCGCTCGTC	GTTATCAGCA	GGCGGCAGAG	240
TGCGATGGCT	GCCATCTGTG	CCTGAACGGC	AGCGATGTAT	TGATCCTCTG	GTGGCCGCTG	300
CCGTCGGATC	CCGGCAGTTA	TCCGCAGGTG	ATCGAACGTT	TGTTTGAACCT	GGCGGGAATG	360
ACGTTGCCGT	CGCTATCCAT	AGCACCGACG	GCGCGTCCGC	AGACAGGGAA	CGGACGCGCC	420
CGATCATTA	GATAAAGGCG	GCTTTTTTTA	TTGCAAAACG	GTAACGGTGA	GGAACCGTTT	480
CACCGTCGGC	GTCACCTCAGT	AACAAGTATC	CATCATGATG	CCTACATCGG	GATCGGCGTG	540
GGCATCCGTT	GCAGATACTT	TTGCGAACAC	CTGACATGAA	TGAGGAAACG	AAATTATGCA	600
AATTACGATC	AAAGCGCACA	TCGGCGGTGA	TTTGGGCGTC	TCCGGTCTGG	GGCTGGGTGC	660
TCAGGGACTG	AAAGGACTGA	ATTCCGCGGC	TTCATCGCTG	GGTTCAGCG	TGGATAAACT	720
GAGCAGCACC	ATCGATAAGT	TGACCTCCGC	GCTGACTTCG	ATGATGTTTG	GCGGCGCGCT	780
GGCGCAGGGG	CTGGGCGCCA	GCTCGAAGGG	GCTGGGGATG	AGCAATCAAC	TGGGCCAGTC	840
TTTCGGCAAT	GGCGCGCAGG	GTGCGAGCAA	CCTGCTATCC	GTACCGAAAT	CCGGCGGCGA	900
TGCGTTGTCA	AAAATGTTTG	ATAAAGCGCT	GGACGATCTG	CTGGGTCATG	ACACCGTGAC	960
CAAGCTGACT	AACCAGAGCA	ACCAACTGGC	TAATTCAATG	CTGAACGCCA	GCCAGATGAC	1020
CCAGGGTAAT	ATGAATGCGT	TCGGCAGCGG	TGTGAACAAC	GCACTGTCGT	CCATTCTCGG	1080
CAACGGTCTC	GGCCAGTCGA	TGAGTGGCTT	CTCTCAGCCT	TCTCTGGGGG	CAGGCGGCTT	1140
GCAGGGCCTG	AGCGGCGCGG	GTGCATTCAA	CCAGTTGGGT	AATGCCATCG	GCATGGGCGT	1200

- 63 -

```

GGGGCAGAAT GCTGCGCTGA GTGCGTTGAG TAACGTCAGC ACCCACGTAG ACGGTAACAA 1260
CCGCCACTTT GTAGATAAAG AAGATCGCGG CATGGCGAAA GAGATCGGCC AGTTTATGGA 1320
TCAGTATCCG GAAATATTTCG GTAAACCGGA ATACCAGAAA GATGGCTGGA GTTCGCCGAA 1380
GACGGACGAC AAATCCTGGG CTAAAGCGCT GAGTAAACCG GATGATGACG GTATGACCGG 1440
CGCCAGCATG GACAAATTCC GTCAGGCGAT GGGTATGATC AAAAGCGCGG TGGCGGGTGA 1500
TACCGGCAAT ACCAACCTGA ACCTGCGTGG CGCGGGCGGT GCATCGCTGG GTATCGATGC 1560
GGCTGTCGTC GGCGATAAAA TAGCCAACAT GTCGCTGGGT AAGCTGGCCA ACGCCTGATA 1620
ATCTGTGCTG GCCTGATAAA GCGGAAACGA AAAAGAGAC GGGGAAGCCT GTCTCTTTTC 1680
TTATTATGCG GTTATGCGG TTACCTGGAC CGGTTAATCA TCGTCATCGA TCTGGTACAA 1740
ACGCACATTT TCCCGTTCAT TCGCGTCGTT ACGCGCCACA ATCGCGATGG CATCTTCCTC 1800
GTCGCTCAGA TTGCGCGGCT GATGGGGAAC GCCGGGTGGA ATATAGAGAA ACTCGCCGGC 1860
CAGATGGAGA CACGTCTGCG ATAAATCTGT GCCGTAACGT GTTCTATCC GCCCCTTTAG 1920
CAGATAGATT GCGGTTTCGT AATCAACATG GTAATGCGGT TCCGCCTGTG CGCCGGCCGG 1980
GATCACCACA ATATTCATAG AAAGCTGTCT TGCACCTACC GTATCGCGGG AGATACCGAC 2040
AAAATAGGGC AGTTTTTGCG TGGTATCCGT GGGGTGTTCC GGCCTGACAA TCTTGAGTTG 2100
GTTGTCATC ATCTTCTCC ATCTGGGCGA CCTGATCGGT T 2141

```

(2) INFORMATION FOR SEQ ID NO:3:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 403 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS:
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

```

Met Ser Leu Asn Thr Ser Gly Leu Gly Ala Ser Thr Met Gln Ile Ser
1           5           10           15
Ile Gly Gly Ala Gly Gly Asn Asn Gly Leu Leu Gly Thr Ser Arg Gln
20           25           30
Asn Ala Gly Leu Gly Gly Asn Ser Ala Leu Gly Leu Gly Gly Gly Asn
35           40           45
Gln Asn Asp Thr Val Asn Gln Leu Ala Gly Leu Leu Thr Gly Met Met
50           55           60
Met Met Met Ser Met Met Gly Gly Gly Gly Leu Met Gly Gly Gly Leu
65           70           75           80

```

- 64 -

[illegible]

(2) INFORMATION FOR SEQ ID NO:4:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1288 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

AAGCTTCGGC ATGGCACGTT TGACCGTTGG GTCGGCAGGG TACGTTTGAA TTATTCATAA	60
GAGGAATACG TTATGAGTCT GAATACAAGT GGGCTGGGAG CGTCAACGAT GCAAATTTCT	120
ATCGGCGGTG CGGGCGGAAA TAACGGGTTG CTGGGTACCA GTCGCCAGAA TGCTGGGTTG	180
GGTGGCAATT CTGCACTGGG GCTGGGCGGC GGTAAATCAA ATGATACCGT CAATCAGCTG	240
GCTGGCTTAC TCACCGGCAT GATGATGATG ATGAGCATGA TGGGCGGTGG TGGGCTGATG	300
GGCGGTGGCT TAGGCGGTGG CTTAGGTAAT GGCTTGGGTG GCTCAGGTGG CCTGGGCGAA	360
GGACTGTCTGA ACGCGCTGAA CGATATGTTA GCGGTTTCGC TGAACACGCT GGGCTCGAAA	420
GGCGGCAACA ATACCAC TTC AACAAACAAT TCCCCGCTGG ACCAGGCGCT GGGTATTAAC	480
TCAACGTCCC AAAACGACGA TTCCACCTCC GGCACAGATT CCACCTCAGA CTCCAGCGAC	540
CCGATGCAGC AGCTGCTGAA GATGTTTCAGC GAGATAATGC AAAGCCTGTT TGGTGATGGG	600
CAAGATGGCA CCCAGGGCAG TTCCTCTGGG GGCAAGCAGC CGACCGAAGG CGAGCAGAAC	660
GCCTATAAAA AAGGAGTCAC TGATGCGCTG TCGGGCCTGA TGGGTAATGG TCTGAGCCAG	720
CTCCTTGSCA ACGGGGGACT GGGAGGTGGT CAGGGCGGTA ATGCTGGCAC GGGTCTTGAC	780
GGTTCGTCGC TGGGCGGCAA AGGGCTGCAA AACCTGAGCG GGCCGGTGGA CTACCAGCAG	840
TTAGGTAACG CCGTGGGTAC CGGTATCGGT ATGAAAGCGG GCATTTCAGGC GCTGAATGAT	900
ATCGGTACGC ACAGGCACAG TTCAACCCGT TCTTTCGTCA ATAAAGGCGA TCGGGCGATG	960
GCGAAGGAAA TCGGTCAGTT CATGGACCAG TATCCTGAGG TGTTTGGCAA GCCGCAGTAC	1020
CAGAAAGGCC CGGGTCAGGA GGTGAAAACC GATGACAAAT CATGGGCAAA AGCACTGAGC	1080
AAGCCAGATG ACGACGGAAT GACACCAGCC AGTATGGAGC AGTTCAACAA AGCCAAGGGC	1140
ATGATCAAAA GGCCCATGGC GGGTGATACC GGCAACGGCA ACCTGCAGGC ACGCGGTGCC	1200
GGTGGTTCTT CGCTGGGTAT TGATGCCATG ATGGCCGGTG ATGCCATTAA CAATATGGCA	1260
CTTGGCAAGC TGGGCGCGGC TTAAGCTT	1288

- 66 -

(2) INFORMATION FOR SEQ ID NO:5:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 341 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS:
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

```

Met Gln Ser Leu Ser Leu Asn Ser Ser Ser Leu Gln Thr Pro Ala Met
 1           5           10           15

Ala Leu Val Leu Val Arg Pro Glu Ala Glu Thr Thr Gly Ser Thr Ser
 20           25           30

Ser Lys Ala Leu Gln Glu Val Val Val Lys Leu Ala Glu Glu Leu Met
 35           40           45

Arg Asn Gly Gln Leu Asp Asp Ser Ser Pro Leu Gly Lys Leu Leu Ala
 50           55           60

Lys Ser Met Ala Ala Asp Gly Lys Ala Gly Gly Gly Ile Glu Asp Val
 65           70           75           80

Ile Ala Ala Leu Asp Lys Leu Ile His Glu Lys Leu Gly Asp Asn Phe
 85           90           95

Gly Ala Ser Ala Asp Ser Ala Ser Gly Thr Gly Gln Gln Asp Leu Met
100           105           110

Thr Gln Val Leu Asn Gly Leu Ala Lys Ser Met Leu Asp Asp Leu Leu
115           120           125

Thr Lys Gln Asp Gly Gly Thr Ser Phe Ser Glu Asp Asp Met Pro Met
130           135           140

Leu Asn Lys Ile Ala Gln Phe Met Asp Asp Asn Pro Ala Gln Phe Pro
145           150           155           160

Lys Pro Asp Ser Gly Ser Trp Val Asn Glu Leu Lys Glu Asp Asn Phe
165           170           175

Leu Asp Gly Asp Glu Thr Ala Ala Phe Arg Ser Ala Leu Asp Ile Ile
180           185           190

Gly Gln Gln Leu Gly Asn Gln Gln Ser Asp Ala Gly Ser Leu Ala Gly
195           200           205

Thr Gly Gly Gly Leu Gly Thr Pro Ser Ser Phe Ser Asn Asn Ser Ser
210           215           220

Val Met Gly Asp Pro Leu Ile Asp Ala Asn Thr Gly Pro Gly Asp Ser
225           230           235           240

Gly Asn Thr Arg Gly Glu Ala Gly Gln Leu Ile Gly Glu Leu Ile Asp
245           250           255

```

- 67 -

Arg Gly Leu Gln Ser Val Leu Ala Gly Gly Gly Leu Gly Thr Pro Val
 260 265 270
 Asn Thr Pro Gln Thr Gly Thr Ser Ala Asn Gly Gly Gln Ser Ala Gln
 275 280 285
 Asp Leu Asp Gln Leu Leu Gly Gly Leu Leu Leu Lys Gly Leu Glu Ala
 290 295 300
 Thr Leu Lys Asp Ala Gly Gln Thr Gly Thr Asp Val Gln Ser Ser Ala
 305 310 315 320
 Ala Gln Ile Ala Thr Leu Leu Val Ser Thr Leu Leu Gln Gly Thr Arg
 325 330 335
 Asn Gln Ala Ala Ala
 340

(2) INFORMATION FOR SEQ ID NO:6:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1026 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

ATGCAGAGTC TCAGTCTTAA CAGCAGCTCG CTGCAAACCC CGGCAATGGC CCTTGTCTCTG	60
GTACGTCCTG AAGCCGAGAC GACTGGCAGT ACGTCGAGCA AGGCGCTTCA GGAAGTTGTC	120
GTGAAGCTGG CCGAGGAACT GATGCGCAAT GGTCAACTCG ACGACAGCTC GCCATTGGGA	180
AAACTGTTGG CCAAGTCGAT GGCCGCAGAT GGCAAGGCGG GCGGCGGTAT TGAGGATGTC	240
ATCGCTGCGC TGGACAAGCT GATCCATGAA AAGCTCGGTG ACAACTTCGG CGCGTCTGCG	300
GACAGCGCCT CGGGTACCGG ACAGCAGGAC CTGATGACTC AGGTGCTCAA TGGCCTGGCC	360
AAGTCGATGC TCGATGATCT TCTGACCAAG CAGGATGGCG GGACAAGCTT CTCCGAAGAC	420
GATATGCCGA TGCTGAACAA GATCGCGCAG TTCATGGATG ACAATCCCGC ACAGTTTCCC	480
AAGCCGGACT CGGGCTCCTG GGTGAACGAA CTCAAGGAAG ACAACTTCCT TGATGGCGAC	540
GAAACGGCTG CGTTCCGTTT GGCACGCGAC ATCATTGGCC AGCAACTGGG TAATCAGCAG	600
AGTGACGCTG GCACTCTGGC AGGGACGGGT GGAGGTCCTG GCACTCCGAG CAGTTTTTCC	660
AACAACTCGT CCGTGATGGG TGATCCGCTG ATCGACGCCA ATACCGGTCC CGGTGACAGC	720
GGCAATACCC GTGGTGAAGC GGGGCAACTG ATCGGCGAGC TTATCGACCG TGGCCTGCAA	780
TCGGTATTGG CCGGTGGTGG ACTGGGCACA CCCGTAAACA CCCCAGAC CCGTACGTCG	840

- 68 -

GCGAATGGCG GACAGTCCGC TCAGGATCTT GATCAGTTGC TGGGCGGCTT GCTGCTCAAG 900
 GGCCTGGAGG CAACGCTCAA GGATGCCGGG CAAACAGGCA CCGACGTGCA GTCGAGCGCT 960
 GCGCAAATCG CCACCTTGCT GGTCAGTACG CTGCTGCAAG GCACCCGCAA TCAGGCTGCA 1020
 GCCTGA 1026

(2) INFORMATION FOR SEQ ID NO:7:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 344 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS:
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:

Met Ser Val Gly Asn Ile Gln Ser Pro Ser Asn Leu Pro Gly Leu Gln
 1 5 10 15
 Asn Leu Asn Leu Asn Thr Asn Thr Asn Ser Gln Gln Ser Gly Gln Ser
 20 25 30
 Val Gln Asp Leu Ile Lys Gln Val Glu Lys Asp Ile Leu Asn Ile Ile
 35 40 45
 Ala Ala Leu Val Gln Lys Ala Ala Gln Ser Ala Gly Gly Asn Thr Gly
 50 55 60
 Asn Thr Gly Asn Ala Pro Ala Lys Asp Gly Asn Ala Asn Ala Gly Ala
 65 70 75 80
 Asn Asp Pro Ser Lys Asn Asp Pro Ser Lys Ser Gln Ala Pro Gln Ser
 85 90 95
 Ala Asn Lys Thr Gly Asn Val Asp Asp Ala Asn Asn Gln Asp Pro Met
 100 105 110
 Gln Ala Leu Met Gln Leu Leu Glu Asp Leu Val Lys Leu Leu Lys Ala
 115 120 125
 Ala Leu His Met Gln Gln Pro Gly Gly Asn Asp Lys Gly Asn Gly Val
 130 135 140
 Gly Gly Ala Asn Gly Ala Lys Gly Ala Gly Gly Gln Gly Gly Leu Ala
 145 150 155 160
 Glu Ala Leu Gln Glu Ile Glu Gln Ile Leu Ala Gln Leu Gly Gly Gly
 165 170 175
 Gly Ala Gly Ala Gly Gly Ala Gly Gly Gly Val Gly Gly Ala Gly Gly
 180 185 190
 Ala Asp Gly Gly Ser Gly Ala Gly Gly Ala Gly Gly Ala Asn Gly Ala
 195 200 205

- 69 -

Asp Gly Gly Asn Gly Val Asn Gly Asn Gln Ala Asn Gly Pro Gln Asn
 210 215 220
 Ala Gly Asp Val Asn Gly Ala Asn Gly Ala Asp Asp Gly Ser Glu Asp
 225 230 235 240
 Gln Gly Gly Leu Thr Gly Val Leu Gln Lys Leu Met Lys Ile Leu Asn
 245 250 255
 Ala Leu Val Gln Met Met Gln Gln Gly Gly Leu Gly Gly Gly Asn Gln
 260 265 270
 Ala Gln Gly Gly Ser Lys Gly Ala Gly Asn Ala Ser Pro Ala Ser Gly
 275 280 285
 Ala Asn Pro Gly Ala Asn Gln Pro Gly Ser Ala Asp Asp Gln Ser Ser
 290 295 300
 Gly Gln Asn Asn Leu Gln Ser Gln Ile Met Asp Val Val Lys Glu Val
 305 310 315 320
 Val Gln Ile Leu Gln Gln Met Leu Ala Ala Gln Asn Gly Gly Ser Gln
 325 330 335
 Gln Ser Thr Ser Thr Gln Pro Met
 340

(2) INFORMATION FOR SEQ ID NO:8:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1035 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:8:

ATGTCAGTCG GAAACATCCA GAGCCCGTCG AACCTCCCGG GTCTGCAGAA CCTGAACCTC	60
AACACCAACA CCAACAGCCA GCAATCGGGC CAGTCCGTGC AAGACCTGAT CAAGCAGGTC	120
GAGAAGGACA TCCTCAACAT CATCGCAGCC CTCGTGCAGA AGGCCGCACA GTCGGCGGGC	180
GGCAACACCG GTAACACCGG CAACGCGCCG GCGAAGGACG GCAATGCCAA CGCGGGCGCC	240
AACGACCCGA GCAAGAACGA CCCGAGCAAG AGCCAGGCTC CGCAGTCGGC CAACAAGACC	300
GGCAACGTCG ACGACGCCAA CAACCAGGAT CCGATGCAAG CGCTGATGCA GCTGCTGGAA	360
GACCTGGTGA AGCTGCTGAA GCGGGCCCTG CACATGCAGC AGCCCGGCGG CAATGACAAG	420
GGCAACGGCG TGGGCGGTGC CAACGGCGCC AAGGGTGCCG GCGGCCAGGG CGGCCTGGCC	480
GAAGCGCTGC AGGAGATCGA GCAGATCCTC GCCCAGCTCG GCGGCGGCGG TGCTGGCGCC	540
GGCGGCGCGG GTGGCGGTGT CGGCGGTGCT GGTGGCGCGG ATGGCGGCTC CGGTGCGGGT	600

- 70 -

```

GGCGCAGGCG GTGCGAACGG CGCCGACGGC GGCAATGGCG TGAACGGCAA CCAGGCGAAC      660
GGCCCGCAGA ACGCAGGCGA TGTC AACGGT GCCAACGGCG CGGATGACGG CAGCGAAGAC      720
CAGGGCGGCC TCACCGGCGT GCTGCAAAAG CTGATGAAGA TCCTGAACGC GCTGGTGCAG      780
ATGATGCAGC AAGGCGGCCT CGGCGGCGGC AACCAGGCGC AGGGCGGCTC GAAGGGTGCC      840
GGCAACGCCT CGCCGGCTTC CGGCGCGAAC CCGGGCGCGA ACCAGCCCGG TTCGGCGGAT      900
GATCAATCGT CCGGCCAGAA CAATCTGCAA TCCCAGATCA TGGATGTGGT GAAGGAGGTC      960
GTCCAGATCC TG CAGCAGAT GCTGGCGGCG CAGAACGGCG GCAGCCAGCA GTCCACCTCG     1020
ACGCAGCCGA TGTA A                                     1035

```

(2) INFORMATION FOR SEQ ID NO:9:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 26 amino acids
 - (B) TYPE: amino acid
 - (C) STRANDEDNESS:
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:9:

```

Thr Leu Ile Glu Leu Met Ile Val Val Ala Ile Ile Ala Ile Leu Ala
1           5           10           15
Ala Ile Ala Leu Pro Ala Tyr Gln Asp Tyr
                20           25

```

(2) INFORMATION FOR SEQ ID NO:10:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 20 amino acids
 - (B) TYPE: amino acid
 - (C) STRANDEDNESS:
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:10:

```

Ser Ser Gln Gln Ser Pro Ser Ala Gly Ser Glu Gln Gln Leu Asp Gln
1           5           10           15
Leu Leu Ala Met
                20

```

- 71 -

WHAT IS CLAIMED:

1. A method of producing plant seeds which impart pathogen resistance to plants grown from the seeds, said method comprising:
- 5 applying a hypersensitive response elicitor polypeptide or protein in a non-infectious form to a plant seed under conditions effective to impart pathogen resistance to a plant grown from the seeds.
- 10 2. A method according to claim 1, wherein the hypersensitive response elicitor polypeptide or protein is in isolated form.
- 15 3. A method according to claim 2, wherein the hypersensitive response elicitor polypeptide or protein corresponds to that derived from a pathogen selected from the group consisting of *Erwinia*, *Pseudomonas*, *Xanthomonas*, *Phytophthora*, and mixtures thereof.
- 20 4. A method according to claim 3, wherein the hypersensitive response elicitor polypeptide or protein corresponds to that derived from *Erwinia chrysanthemi*.
- 25 5. A method according to claim 3, wherein the hypersensitive response elicitor polypeptide or protein corresponds to that derived from *Erwinia amylovora*.
- 30 6. A method according to claim 3, wherein the hypersensitive response elicitor polypeptide or protein corresponds to that derived from *Pseudomonas syringae*.
7. A method according to claim 3, wherein the hypersensitive response elicitor polypeptide or protein

- 72 -

corresponds to that derived from *Pseudomonas solanacearum*.

8. A method according to claim 3, wherein the
5 hypersensitive response elicitor polypeptide or protein
corresponds to that derived from *Xanthomonas campestris*.

9. A method according to claim 3, wherein the
hypersensitive response elicitor polypeptide or protein
10 corresponds to a *Phytophthora* species.

10. A method according to claim 2, wherein the
plant is selected from the group consisting of dicots and
monocots.

11. A method according to claim 10, wherein
the plant is selected from the group consisting of rice,
wheat, barley, rye, oats, cotton, sunflower, canola,
peanut, corn, potato, sweet potato, bean, pea, chicory,
20 lettuce, endive, cabbage, cauliflower, broccoli, turnip,
radish, spinach, onion, garlic, eggplant, pepper, celery,
carrot, squash, pumpkin, zucchini, cucumber, apple, pear,
melon, strawberry, grape, raspberry, pineapple, soybean,
tobacco, tomato, sorghum, and sugarcane.

12. A method according to claim 10, wherein
the plant is selected from the group consisting of rose,
Saintpaulia, petunia, *Pelargonium*, poinsettia,
chrysanthemum, carnation, and zinnia.

13. A method according to claim 2, wherein the
pathogen to which the plant is resistant is selected from
the group consisting of viruses, bacteria, fungi, and
combinations thereof.

35

- 73 -

14. A method according to claim 2, wherein said applying is carried out by spraying, injection, coating, dusting or immersion.

5 15. A method according to claim 2, wherein the hypersensitive response elicitor polypeptide or protein is applied to plant seeds as a composition further comprising a carrier.

10 16. A method according to claim 15, wherein the carrier is selected from the group consisting of water, aqueous solutions, slurries, and powders.

15 17. A method according to claim 15, wherein the composition contains greater than .5 nM of the hypersensitive response elicitor polypeptide or protein.

20 18. A method according to claim 15, wherein the composition further contains additives selected from the group consisting of fertilizer, insecticide, nematicide, fungicide, herbicide, and mixtures thereof.

25 19. A method according to claim 1, wherein the hypersensitive response elicitor polypeptide or protein is applied as bacteria which do not cause disease and are transformed with a gene encoding the hypersensitive response elicitor polypeptide or protein.

30 20. A method according to claim 1, wherein the hypersensitive response elicitor polypeptide or protein is applied as bacteria which cause disease in some plant species, but not in those whose seeds are subjected to said applying, and contain a gene encoding the hypersensitive response elicitor polypeptide or protein.

35

- 74 -

21. A method according to claim 2, wherein said applying causes infiltration of the polypeptide or protein into the plant seed.

5 22. A method according to claim 2 further comprising:

 planting in soil the seeds to which the hypersensitive response elicitor protein or polypeptide has been applied and

10 propagating plants from the planted seeds.

23. A method according to claim 22 further comprising:

15 applying the hypersensitive response elicitor polypeptide or protein to the propagated plants to enhance the plant's pathogen resistance.

24. A method according to claim 2, wherein the hypersensitive response elicitor protein or polypeptide
20 is a fungal hypersensitive response elicitor.

25 25. A pathogen-resistance imparting plant seed to which a non-infectious hypersensitive response elicitor polypeptide or protein has been applied.

26. A pathogen-resistance imparting plant seed according to claim 25, wherein the hypersensitive response elicitor polypeptide or protein is in isolated form.
30

27. A pathogen-resistance imparting plant seed according to claim 26, wherein the hypersensitive response elicitor polypeptide or protein corresponds to that derived from a pathogen selected from the group

- 75 -

consisting of *Erwinia*, *Pseudomonas*, *Xanthomonas*,
Phytophthora, and mixtures thereof.

28. A pathogen-resistance imparting plant seed
5 according to claim 27, wherein the hypersensitive
response elicitor polypeptide or protein corresponds to
that derived from *Erwinia chrysanthemi*.

29. A pathogen-resistance imparting plant seed
10 according to claim 27, wherein the hypersensitive
response elicitor polypeptide or protein corresponds to
that derived from *Erwinia amylovora*.

30. A pathogen-resistance imparting plant seed
15 according to claim 27, wherein the hypersensitive
response elicitor polypeptide or protein corresponds to
that derived from *Pseudomonas syringae*.

31. A pathogen-resistance imparting plant seed
20 according to claim 27, wherein the hypersensitive
response elicitor polypeptide or protein corresponds to
that derived from *Pseudomonas solanacearum*.

32. A pathogen-resistance imparting plant seed
25 according to claim 27, wherein the hypersensitive
response elicitor polypeptide or protein corresponds to
that derived from *Xanthomonas campestris*.

33. A pathogen-resistance imparting plant seed
30 according to claim 27, wherein the hypersensitive
response polypeptide or protein corresponds to that
derived from a *Phytophthora* species.

34. A pathogen-resistance imparting plant seed
35 according to claim 26, wherein the plant seed is for

- 76 -

plants selected from the group consisting of dicots and monocots.

35. A pathogen-resistance imparting plant seed
5 according to claim 34, wherein the plant is selected from
the group consisting of rice, wheat, barley, rye, oats,
cotton, sunflower, canola, peanut, potato, sweet potato,
bean, pea, chicory, lettuce, endive, cabbage,
10 cauliflower, broccoli, turnip, radish, spinach, onion,
garlic, eggplant, pepper, celery, carrot, squash,
pumpkin, zucchini, cucumber, apple, pear, melon,
strawberry, grape, raspberry, pineapple, soybean,
tobacco, tomato, sorghum, and sugarcane.

15 36. A pathogen-resistance imparting plant seed
according to claim 34, wherein the plant is selected from
the group consisting of rose, *Saintpaulia*, petunia,
Pelargonium, poinsettia, chrysanthemum, carnation, and
zinnia.

20 37. A pathogen-resistance imparting plant seed
according to claim 27, wherein the pathogen to which the
plant is resistant is selected from the group consisting
of a virus, bacterium, fungus, nematode, and combinations
25 thereof.

38. A pathogen-resistance imparting plant seed
according to claim 25, wherein the plant seed cells are
in contact with bacteria which do not cause disease and
30 are transformed with a gene encoding the hypersensitive
response elicitor polypeptide or protein.

39. A pathogen-resistance imparting plant seed
according to claim 25, wherein the plant seed cells are
35 in contact with bacteria which do not cause disease in

- 77 -

the plant, but do cause disease in other plant species, and contain a gene encoding the hypersensitive response elicitor polypeptide or protein.

5 40. A pathogen-resistance imparting plant seed according to claim 26, wherein the plant seed is infiltrated with the polypeptide or protein.

10 41. A method of imparting pathogen resistance to plants comprising:
 providing a transgenic plant seed transformed with a DNA molecule encoding a hypersensitive response elicitor polypeptide or protein;
 planting the transgenic plant seed in
15 soil; and
 propagating a plant from the planted seed under conditions effective to impart pathogen resistance to the plant.

20 42. A method according to claim 39, wherein the hypersensitive response elicitor polypeptide or protein corresponds to that derived from a pathogen selected from the group consisting of *Erwinia*, *Pseudomonas*, *Xanthomonas*, *Phytophthora*, and mixtures
25 thereof.

 43. A method according to claim 42, wherein the hypersensitive response elicitor polypeptide or protein corresponds to that derived from *Erwinia*
30 *chrysanthemi*.

 44. A method according to claim 42, wherein the hypersensitive response elicitor polypeptide or protein corresponds to that derived from *Erwinia*
35 *amylovora*.

- 78 -

45. A method according to claim 42, wherein the hypersensitive response elicitor polypeptide or protein corresponds to that derived from *Pseudomonas syringae*.

5

46. A method according to claim 42, wherein the hypersensitive response elicitor polypeptide or protein corresponds to that derived from *Pseudomonas solanacearum*.

10

47. A method according to claim 42, wherein the hypersensitive response elicitor polypeptide or protein corresponds to that derived from *Xanthomonas campestris*.

15

48. A method according to claim 42, wherein the hypersensitive response elicitor polypeptide or protein corresponds to that derived from a *Phytophthora* species.

20

49. A method according to claim 41, wherein the plant is selected from the group consisting of dicots and monocots.

25

50. A method according to claim 49, wherein the plant is selected from the group consisting of rice, wheat, barley, rye, oats, cotton, sunflower, canola, peanut, corn, potato, sweet potato, bean, pea, chicory, lettuce, endive, cabbage, cauliflower, broccoli, turnip, radish, spinach, onion, garlic, eggplant, pepper, celery, carrot, squash, pumpkin, zucchini, cucumber, apple, pear, melon, strawberry, grape, raspberry, pineapple, soybean, tobacco, tomato, sorghum, and sugarcane.

30

- 79 -

51. A method according to claim 49, wherein the plant is selected from the group consisting of rose, *Saintpaulia*, petunia, *Pelargonium*, poinsettia, chrysanthemum, carnation, and zinnia.

5

52. A method according to claim 41, wherein the pathogen to which the plant is resistant is selected from the group consisting of viruses, bacteria, fungi, and combinations thereof.

10

53. A method according to claim 41 further comprising:

applying the hypersensitive response elicitor polypeptide or protein to the propagated plants to enhance the plant's pathogen resistance.

15

54. A method according to claim 41, wherein the hypersensitive response elicitor protein or polypeptide is a fungal hypersensitive response elicitor.

20

55. A plant produced by the method of claim 22.

56. A plant seed from the plant produced by the method of claim 22.

25

57. A plant propagule from the plant produced by the method of claim 22.

58. A plant produced by the method of claim 41.

30

59. A plant seed from the plant produced by the method of claim 41.

35

- 80 -

60. A plant propagule from the plant produced
by the method of claim 41.

INTERNATIONAL SEARCH REPORT

International application No.

PCT/US97/22629

A. CLASSIFICATION OF SUBJECT MATTER

IPC(6) : Please See Extra Sheet.

US CL : 800/200, 250; 514/2; 530/370

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

U.S. : 800/200, 250; 514/2; 530/370

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

Sequence, CAS - Agriculture and Bioscience Clusters

Search terms: hypersensitive, elicitor, harpin, seed, spore, tuber, kernel, pit, pip

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X — Y	US 5,550,228 A (GODIARD, et al.) 27 August 1996, entire document but specifically col. 4, lines 46-59, col. 5, lines 30-41, cols. 5-6 lines 59-7.	19-20, 41-42, 44, 46, 49-50, 52, 58- 60 43, 45, 47-49, 51
A	US 5,552,527 A (GODIARD et al.) 09 September 1996, entire document.	
Y	WO 94/01546 A1 (CORNELL RESEARCH FOUNDATION, INC) 20 January 1994, entire document.	43, 45, 47-49, 51
A	WO 94/26782 A1 (CORNELL RESEARCH FOUNDATION, INC.) 24 November 1994.	

☒ Further documents are listed in the continuation of Box C. ☐ See patent family annex.

* Special categories of cited documents:	*T* later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
A document defining the general state of the art which is not considered to be of particular relevance	*X* document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
B earlier document published on or after the international filing date	*Y* document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art
L document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)	*A* document member of the same patent family
O document referring to an oral disclosure, use, exhibition or other means	
P document published prior to the international filing date but later than the priority date claimed	

Date of the actual completion of the international search

12 MARCH 1998

Date of mailing of the international search report

07 APR 1998

Name and mailing address of the ISA/US
Commissioner of Patents and Trademarks
Box PCT
Washington, D.C. 20231

Facsimile No. (703) 305-3230

Authorized officer

JENNIFER FARLE

Telephone No. (703) 305-0196

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US97/22629

C (Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X,P	WO 96/39802 A1 (CCORNELL RESEARCH FOUNDATION, INC.) 19 December 1996, entire document.	1-40, 55-57
X,P	WEI, et al., Hypersensitive response induced resistance in plants ,Cornell Research Foundation, Inc.1997.AN ; 639888.	1-40, 55-57
X,P ---- Y,P	QUI, et al., Treatment of Tomato Seed with Harpin Enhances Germination and Growth and Induces Resistance to Ralstonia Solanacearum, Phytopathology. 1997. Vol. 87. No. 6. S80.	1-3, 5, 10-11, 13-16, 21-23, 25-27, 29, 34-35, 37, 40, 55-57 4, 6-9, 12, 17, 18, 24, 28, 30-33, 36
P,A	US 5,708,139 A (COLLMER, et al) 13 January 1998, entire document.	

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US97/22629

A. CLASSIFICATION OF SUBJECT MATTER:

IPC (6):

A01G 1/00, A01H 1/00, 1/02, 5/00, 7/00, 9/00, 11/00, 13/00, 15/00, 17/00; A01N 37/18; A61K 35/78, 35/80, 38/00;
C07K 1/00, 14/00, 16/00, 17/00

(19) World Intellectual Property Organization
International Bureau



(43) International Publication Date
1 March 2001 (01.03.2001)

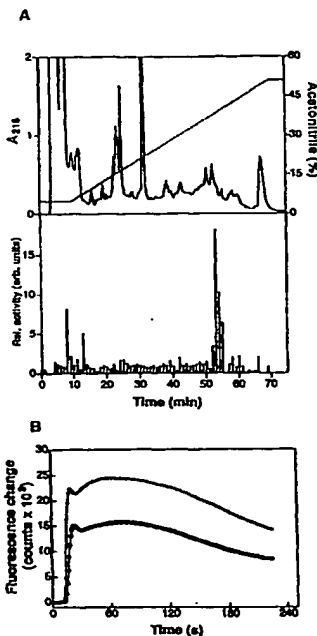
PCT

(10) International Publication Number
WO 01/14888 A1

- (51) International Patent Classification⁷: G01N 33/74, (72) Inventors: CIVELLI, Olivier; 10 Dickens Street, Irvine, CA 92612 (US). NOTHACKER, Hans-Peter; 2003 C. Los Trancos Drive #C, Irvine, CA 92612 (US).
C07K 14/72, 14/575, 14/705
- (21) International Application Number: PCT/US00/21171 (74) Agent: BERLINER, Robert; Fulbright & Jaworski L.L.P., 29th Floor, 865 South Figueroa Street, Los Angeles, CA 90017-2576 (US).
- (22) International Filing Date: 3 August 2000 (03.08.2000)
- (25) Filing Language: English (81) Designated States (*national*): AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CR, CU, CZ, DE, DK, DM, DZ, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, UZ, VN, YU, ZA, ZW.
- (26) Publication Language: English
- (30) Priority Data: 09/379,968 24 August 1999 (24.08.1999) US (84) Designated States (*regional*): ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European

[Continued on next page]

(54) Title: METHOD OF IDENTIFYING UROTENSIN II RECEPTOR ANTAGONISTS



(57) Abstract: The invention provides a method of identifying a urotensin II (U II) receptor agonist or antagonist. The method consists of contacting an isolated U II receptor with one or more candidate compounds under conditions wherein the U II receptor produces a predetermined signal in response to U II, and identifying a candidate compound that alters production of the signal. Such a compound is characterized as a U II receptor agonist or antagonist. Also provided is a method of identifying a U II receptor ligand. The method consists of contacting an isolated U II receptor with one or more candidate compounds in the presence of detectably labeled U II, and identifying a compound that decreases binding of the detectably labeled U II to the U II receptor. Such a compound is characterized as a U II receptor ligand. The invention also provides a composition, having an isolated U II receptor and detectably labeled U II.

WO 01/14888 A1



patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

— *Before the expiration of the time limit for amending the claims and to be republished in the event of receipt of amendments.*

Published:

— *With international search report.*

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

METHOD OF IDENTIFYING UROTENSIN II RECEPTOR ANTAGONISTS

BACKGROUND OF THE INVENTION

The present invention relates generally to the field of medicine and, more specifically, to methods and compositions related to the Urotensin II receptor.

The regulation of vascular smooth muscle tone controls the flow of blood throughout the body and the maintenance of appropriate blood pressure. A small, cyclic peptide called Urotensin II (U II), first identified as a neuropeptide in fish, has been demonstrated to act as a vasoconstrictor in mammalian major arteries.

In view of the role of Urotensin II in regulating the vessels of the circulatory system, it would be beneficial to develop compounds that mimic or antagonize U II activity. These compounds would be valuable as therapeutics in conditions in which abnormal circulatory function plays a role. For example, compounds that mimic the vasoconstrictive action of U II could be used as therapeutics in the prevention and treatment of hypotension and shock, which are often consequences of trauma or hemorrhage. In contrast, compounds that antagonize the vasoconstrictive action of U II could be used as vasodilators to prevent or treat diseases such as angina pectoris, hypertension, myocardial infarction, stroke, and congestive heart failure.

However, the cell surface receptor that binds Urotensin II, and the signal transduction pathway initiated by receptor binding, have not yet been

identified. Therefore, it has not been possible to develop rapid and reliable methods of screening for therapeutic compounds that can be used to regulate or alter U II-mediated physiological and pathological functions.

Thus, there exists a need to identify the U II receptor and to develop methods of screening for compounds that bind to U II receptor or mimic or antagonize U II activity. The present invention satisfies this need and provides related advantages as well.

SUMMARY OF THE INVENTION

The invention provides a method of identifying a urotensin II (U II) receptor agonist or antagonist.

15 The method consists of contacting an isolated U II receptor with one or more candidate compounds under conditions wherein the U II receptor produces a predetermined signal in response to U II, and identifying a candidate compound that alters production of the

20 signal. Such a compound is characterized as a U II receptor agonist or antagonist. Also provided is a method of identifying a U II receptor ligand. The method consists of contacting an isolated U II receptor with one or more candidate compounds in the presence of detectably

25 labeled U II, and identifying a compound that decreases binding of the detectably labeled U II to the U II receptor. Such a compound is characterized as a U II receptor ligand. The invention also provides a composition, having an isolated U II receptor and

30 detectably labeled U II.

BRIEF DESCRIPTION OF THE DRAWINGS

Figure 1A shows the nucleotide sequence (SEQ ID NO:1) and Figure 1B shows the deduced amino acid sequence (SEQ ID NO:2) of rat Urotensin II receptor (GPR14 or
5 SENR) (GenBank accession number U23483).

Figure 2A shows the RP-HPLC profile and corresponding activity of fractions of a bovine hypothalamic tissue extract tested for calcium mobilization in GPR14 expressing CHO cells. Figure 2B
10 shows the comparison of calcium kinetics induced by 0.1 nM of human urotensin II (hU II) (closed circles) or 2.5% of fraction 53 (open circles).

Figure 3A shows the effect of human U II and related peptides on calcium influx in GPR14 expressing
15 CHO cells. Figure 3B shows saturation binding of ¹²⁵I-labeled human U II.

Figure 4A shows qualitative RT-PCR analysis of GPR14 expression in cardiovascular tissues. M, molecular weight marker; 1, negative control; 2, thoracic aorta; 3, abdominal aorta; 4, heart; 5, positive control plasmid.
20 Figure 4B shows the contraction of thoracic aorta strips induced by increasing concentrations of hU II.

Figure 5A shows the tissue distribution of human preprourotensin II (pphU II) on a masterblot. E1
25 contains human kidney RNA; B7 contains human spinal cord RNA. Figure 5B shows the RP-HPLC profile of conditioned medium from HEK 293T kidney cells expressing human (pphU II).

DETAILED DESCRIPTION OF THE INVENTION

The present invention relates to the identification of the receptor for urotensin II (U II) and its signal transduction pathway. The invention thus provides methods and compositions for identifying compounds that specifically bind to or modulate signaling through the U II receptor. Such compounds can be used therapeutically to prevent or ameliorate U II receptor-associated conditions, including cardiovascular diseases.

10 The invention provides a method of identifying a U II receptor agonist or antagonist. The method consists of contacting an isolated U II receptor with a candidate compound under conditions wherein the U II receptor produces a predetermined signal in response to U II, and identifying a compound that alters production of the predetermined signal. A compound that alters production of the signal is characterized as a U II receptor agonist or antagonist.

As used herein, the term "U II receptor" refers to a heptahelical membrane-spanning G protein-coupled polypeptide, previously designated GPR14 (Marchese et al. Genomics 29:335-344 (1995)) or SENR (Tal et al., Bioc. Biophys. Res. Comm. 209:752-759 (1995)), which, as disclosed herein, is the endogenous receptor for urotensin II. The term "U II receptor" encompasses native U II receptor polypeptides from all vertebrate species including but not limited to human, non-human primate, rat, mouse, rabbit, bovine, porcine, ovine, canine, feline, avian, reptile, amphibian or fish. The rat U II receptor nucleotide sequence (SEQ ID NO:1) and encoded amino acid sequence (SEQ ID NO:2) are described in Tal et al., supra (1995), and are shown in Figure 1.

The term "U II receptor" also encompasses polypeptides containing minor modifications with respect to a native U II receptor sequence, and fragments of full-length U II receptor, so long as the modified polypeptide or fragment retains one or more of the biological activities of a native U II receptor. As used herein, the "biological activities" of a native U II receptor refers to its ability to selectively bind U II, and the ability to couple to and signal through a G protein in response to U II. The ability to selectively bind U II and the ability to couple to signal through a G protein in response to U II can be determined by the binding and signaling assays disclosed herein.

A modified U II receptor polypeptide can have one or more additions, deletions, or substitutions of natural or non-natural amino acids relative to the native polypeptide sequence, so long as a biological activity of a native U II receptor is retained. A modification to a polypeptide sequence can be, for example, a conservative change, wherein a substituted amino acid has similar structural or chemical properties, e.g., substitution of an apolar amino acid with another apolar amino acid (such as replacement of leucine with isoleucine). A modification can also be a nonconservative change, wherein a substituted amino acid has different but sufficiently similar structural or chemical properties so as to not adversely affect the desired biological activity, e.g., replacement of an amino acid with an uncharged polar R group with an amino acid with an apolar R group (such as replacement of glycine with tryptophan). Further, a minor modification can be the substitution of an L-configuration amino acid with the corresponding D-configuration amino acid with a non-natural amino acid.

In addition, a minor modification can be a chemical or enzymatic modification to the encoded polypeptide, including but not limited to the following: replacement of hydrogen by an alkyl, acyl, or amino
5 group; esterification of a carboxyl group with a suitable alkyl or aryl moiety; alkylation of a hydroxyl group to form an ether derivative; phosphorylation or dephosphorylation of a serine, threonine or tyrosine residue; or N- or O-linked glycosylation.

10 Those skilled in the art can determine whether minor modifications to the native U II sequence are advantageous. Such modifications can be made, for example, to enhance the stability, bioavailability or bioactivity of the U II receptor. A modified U II
15 receptor polypeptide can be prepared, for example, by recombinant methods, by synthetic methods, by post-synthesis chemical or enzymatic methods, or by a combination of these methods, and tested for its U II receptor biological activity using the signaling and
20 binding assays disclosed herein.

The method of identifying a U II receptor agonist, antagonist or ligand is practiced by contacting an "isolated U II receptor" with a candidate compound. As used herein, the term "isolated U II receptor" refers
25 to a U II receptor that is present in a form or composition different from how it naturally exists in tissue. For example, an isolated U II receptor can be present endogenously in an "isolated cell." As used herein, the term "isolated cell" refers to a cell that is
30 substantially purified from non-cellular tissue components, such as connective tissue fibers. Such a cell can be, for example, a primary cell, either freshly purified from non-cellular tissue components, or cultured

for one or many generations. An "isolated cell" can also be, for example, an established cell line. An example of an isolated cell that endogenously expresses U II receptor is a primary cell or established cell line
5 assayed for its expression of U II receptor mRNA or protein, and determined by such criteria to express U II receptor.

In contrast, a tissue preparation with uncharacterized U II binding activity is not considered
10 herein to be an "isolated U II receptor" or to be present in an "isolated cell." Additionally, a U II receptor present in crude fractions, such as crude membrane fractions, of such tissue preparations, is also not considered herein to be an "isolated U II receptor."

15 In the methods of identifying U II receptor agonists, antagonists and ligands described herein, a U II receptor can also be an "isolated U II receptor" in that it is present at higher abundance or higher density than in cells as found in nature, or present in a cell
20 that does not naturally express U II receptor. Thus, an "isolated U II receptor" can be a U II receptor that is recombinantly expressed, either transiently or stably, in a cell or cell extract. An example of a cell that recombinantly expresses an isolated U II receptor is a
25 Chinese Hamster Ovary (CHO) cell transfected with an expression vector containing a nucleic acid sequence encoding the U II receptor sequence designated SEQ ID NO:2.

Recombinant expression is advantageous in
30 providing a higher level of expression of the polypeptide than is found endogenously, and also allows expression in cells or extracts in which the polypeptide is not

normally found. A recombinant nucleic acid expression construct generally will contain a constitutive or inducible promoter of RNA transcription appropriate for the host cell or transcription-translation system, operatively linked to a nucleotide sequence that encodes the polypeptide of interest. The expression construct can be DNA or RNA, and optionally can be contained in a vector, such as a plasmid or viral vector. Based on knowledge of the nucleic acid sequence encoding U II receptor, one skilled in the art can recombinantly express desired levels of a biologically active U II receptor polypeptide using routine laboratory methods, described, for example, in standard molecular biology technical manuals, such as Sambrook et al., Molecular Cloning: A Laboratory Manual, Cold Spring Harbor Laboratory, New York (1992) and Ansubel et al., Current Protocols in Molecular Biology, John Wiley and Sons, Baltimore, MD (1998).

Exemplary host cells that can be used to express recombinant U II receptor include isolated mammalian primary cells; established mammalian cell lines, such as COS, CHO, HeLa, NIH3T3, HEK 293-T and PC12; amphibian cells, such as *Xenopus* embryos and oocytes; and other vertebrate cells. Exemplary host cells also include insect cells (e.g. *Drosophila*), yeast cells (e.g. *S. cerevisiae*, *S. pombe*, or *Pichia pastoris*) and prokaryotic cells (e.g. *E. coli*), engineered to recombinantly express U II receptor. Additionally, recombinant U II receptor can be expressed in extracts that support transcription and translation, such as reticulocyte lysates and wheat germ extracts.

An "isolated U II receptor" can also be a U II receptor present in a cell obtained from a transgenic

animal, such as a transgenic mouse, that has been engineered by known methods to express recombinant U II receptor in some or all tissues.

5 An appropriate assay for establishing that an isolated cell expresses U II receptor can be determined by those skilled in the art. Such an assay can involve, for example, analysis of expression of U II receptor nucleic acid or expression of U II receptor polypeptide
10 by methods known in the art. Assays for determining expression of U II receptor mRNA include, for example, Northern blots, RT-PCR or *in situ* hybridization analysis. Such methods are described, for example, in standard molecular biology manuals such as Sambrook et al., supra,
15 (1992) and Ansubel et al., supra, (1998). Assays for determining expression of U II receptor protein include, for example, immunoblot analysis, immunoprecipitation, immunofluorescence or immunohistochemistry, using antibodies specific for U II receptor.

20 Antibodies specific for U II receptor can be produced, for example, using whole U II receptor or its peptide fragments as an immunogen. The receptor or peptide can be formulated in an immunogenic composition, such as conjugated to a carrier protein or formulated
25 with an adjuvant, and administered to a laboratory animal. Methods of producing polyclonal or monoclonal antibodies, and their use in detecting protein expression, are well known in the art and are described, for example, in Harlow and Lane, Antibodies: A Laboratory
30 Manual, Cold Spring Harbor Laboratory, New York (1989). Methods of preparing fragments of such antibodies with specific binding activity, such as Fab fragments, and methods of preparing recombinant, chimeric or humanized

antibodies directed against a desired peptide sequence, are also well known in the art.

In the methods of identifying ligands, agonists
5 and antagonists of the U II receptor disclosed herein, a suitable "isolated U II receptor" can also be a U II receptor present in a crude or partially purified fraction of an "isolated cell" expressing a U II receptor, as described above, such as a membrane
10 fraction. A membrane fraction of an isolated cell expressing recombinant U II receptor can be prepared, for example, by the method described in Example III, below.

An "isolated U II receptor" useful in the methods of the invention can also be a substantially
15 purified U II receptor, or active fragment thereof having U II binding activity, or G protein coupled signaling activity, or both, depending on the assay. A substantially purified U II receptor can be prepared, for example, by purification of U II receptor from a tissue,
20 cell or cell extract that endogenously or recombinantly expresses U II receptor. Such purification can be accomplished using fractionation methods known in the art including, for example, column chromatography and immunoaffinity procedures. Additionally, a substantially
25 purified U II receptor or active fragment can be prepared by direct chemical synthesis.

If desired in a particular assay, a substantially purified U II receptor or fragment can be incorporated into a membrane composition to ensure
30 correct folding of the U II receptor. Such a composition can include, for example, a lipid bilayer, such as a natural membrane or liposome, for maintaining the native conformation of the U II receptor. Additionally, for

signaling assays, the membrane composition can include cellular fractions or isolated components necessary for transducing and detecting the desired G protein-coupled signal. Such additional components can include, for
5 example, G proteins, GTP, effector molecules and signal indicator molecules.

The method of identifying a U II receptor agonist, antagonist or ligand involves contacting the isolated U II receptor with one or more "candidate
10 compounds." As used herein, the term "candidate compound" refers to any molecule that potentially acts as a U II receptor agonist, antagonist or ligand in the screening methods disclosed herein. A candidate compound can be a naturally occurring macromolecule, such as a
15 peptide, nucleic acid, carbohydrate, lipid, or any combination thereof. A candidate compound also can be a partially or completely synthetic derivative, analog or mimetic of such a macromolecule, or a small organic or inorganic molecule prepared partly or completely by
20 combinatorial chemistry methods. If desired in a particular assay format, a candidate compound can be detectably labeled or attached to a solid support.

Methods for producing pluralities of compounds, including chemical or biological molecules such as simple
25 or complex organic molecules, metal-containing compounds, carbohydrates, peptides, proteins, peptidomimetics, glycoproteins, lipoproteins, nucleic acids, antibodies, and the like, are well known in the art and are described, for example, in Huse, U.S. Patent No.
30 5,264,563; Francis et al., Curr. Opin. Chem. Biol. 2:422-428 (1998); Tietze et al., Curr. Biol., 2:363-371 (1998); Sofia, Mol. Divers. 3:75-94 (1998); Eichler et al., Med. Res. Rev. 15:481-496 (1995); and the like. Libraries

containing large numbers of natural and synthetic compounds also can be obtained from commercial sources.

The number of different candidate compounds to test in the methods of the invention will depend on the application of the method. For example, one or a small number of candidate compounds can be advantageous in manual screening procedures, or when it is desired to compare efficacy among several identified ligands, agonists or antagonists. However, it is generally understood that the larger the number of candidate compounds, the greater the likelihood of identifying a compound having the desired activity in a screening assay. Additionally, large numbers of compounds can be processed in high-throughput automated screening assays. Therefore, "one or more candidate compounds" can contain, for example, greater than about 10^3 different compounds, preferably greater than about 10^5 different compounds, more preferably, greater than about 10^7 different compounds.

20

The method of identifying a U II receptor agonist or antagonist involves contacting the U II receptor with candidate compounds under conditions where U II receptor produces a predetermined signal in response to "U II." As used herein, the term "U II" refers to a natural or synthetic cyclic peptide containing the core structure CFWDYC (SEQ ID NO:3), with a disulfide bridge between the Cys residues at positions 1 and 6 of SEQ ID NO:3, and which activates signaling through the U II receptor at subnanomolar concentrations. Thus, the term "U II" includes peptides having the core structure CFWDYC, and the flanking amino acid residues of a U II from any species, including, for example, teleost fishes and frog (see Conlon et al., J. Reg. Peptides 69:95-103

30

(1997)), as well as human (see Coulaouarn et al., Proc. Natl. Acad. Sci. USA 95:15803-15808 (1998)).

In particular, the term "U II" includes cyclic peptides having the goby U II sequence AGTADCFWDYCV (SEQ ID NO:4), and the human U II sequence ETPDCFWKTCV (SEQ ID NO:5), which activate signaling in an isolated cell expressing a U II receptor with half-maximal effective concentrations of 0.14 nM and 0.10 nM, respectively. The term "U II" also includes peptides with U II activity having the U II core structure and non-naturally occurring N-terminal or C-terminal sequences.

Biologically active U II can be prepared by purification from cells or tissues that express U II, such as kidney or spinal cord, or from cells that recombinantly express preprorenin II and correctly process it into active U II (see Example V, below). More conveniently, U II can be synthesized by conventional methods of peptide synthesis.

The invention provides a method of identifying U II receptor agonists. As used herein, the term "U II receptor agonist" refers to any compound other than U II that promotes or enhances normal signal transduction through U II receptor with high selectivity and high potency. Thus, a U II receptor agonist can promote signaling through U II receptor at high potency, such as half-maximal signaling at concentrations of less than 100 nM, preferably less than 10 nM, most preferably less than 1 nM. Additionally, a U II receptor agonist can promote signaling through U II receptor with high selectivity, such as at concentrations substantially lower, such as at least 100-fold lower or at least 1000-fold lower, than required to promote signaling through

other G protein coupled receptors, such as, for example, somatostatin 2A receptor.

Examples of U II agonists disclosed herein are [I-Tyr⁹] human U II and [¹²⁵I-Tyr⁹] human U II. As indicated in Example III and Table 1, monoiodination of the core Tyr residue of human U II does not preclude its ability to activate signaling through the U II receptor at subnanomolar concentrations, whereas di-iodinated human U II is a less potent agonist. A further example of a U II agonist, of lesser potency, is [D-Trp⁷]Urotensin II-(5-10) (SEQ ID NO:6).

A U II receptor agonist can act by any agonistic mechanism, such as by binding a U II receptor at the normal U II binding site, thereby promoting U II receptor signaling. A U II receptor agonist can also act, for example, by potentiating the binding activity of U II or signaling activity of U II receptor. The methods of the invention can advantageously be used to identify a U II receptor agonist that acts through any agonistic mechanism.

The invention also provides a method of identifying U II receptor antagonists. As used herein, the term "U II receptor antagonist" refers to a compound that selectively inhibits or decreases normal signal transduction through the U II receptor, with high potency and high selectivity. A U II receptor antagonist can act by any antagonistic mechanism, such as by binding U II or U II receptor, thereby inhibiting binding between U II and U II receptor. A U II receptor antagonist can also act indirectly, for example, by modifying U II or U II receptor. The methods of the invention can

advantageously be used to identify a U II receptor antagonist that acts through any antagonistic mechanism.

The method to identify a U II receptor agonist or antagonist is performed under conditions where U II
5 produces a "predetermined signal" in response to U II. As used herein, the term "predetermined signal" refers to a readout, detectable by any analytical means, that is a qualitative or quantitative indication of activation of G
10 protein-dependent signal transduction through U II receptor. The term "G protein" refers to a class of heterotrimeric GTP binding proteins, with subunits designated $G\alpha$, $G\beta$ and $G\gamma$, that couple to seven-transmembrane cell surface receptors to transduce a variety of extracellular stimuli, including light,
15 neurotransmitters, hormones and odorants to various intracellular effector proteins. G proteins are present in both eukaryotic and prokaryotic organisms, including mammals, other vertebrates, *Drosophila* and yeast.

As disclosed herein, contacting an isolated
20 U II receptor with U II leads to activation of Ca^{2+} influx in a signaling composition, such as CHO cells recombinantly expressing the U II receptor. Therefore, an exemplary predetermined signal that is a qualitative or quantitative indication of activation of G protein-
25 dependent signal transduction through U II receptor is Ca^{2+} influx, which can be measured, for example, using the calcium indicator fluo-3 and fluorescence monitoring system described in Example I, below.

If desired, a predetermined signal other than
30 Ca^{2+} influx can be used as the readout in the methods of the invention. The specificity of a G protein for cell-surface receptors is determined by the C-terminal five

amino acids of the G α subunit. The nucleotide sequences and signal transduction pathways of different classes and subclasses of G α subunits in a variety of eukaryotic and prokaryotic organisms are well known in the art. Thus, any convenient G-protein mediated signal transduction pathway can be assayed by preparing a chimeric G α containing the C-terminal residues of a G α that couples to U II receptor, such as G α q, with the remainder of the protein corresponding to a G α that couples to the signal transduction pathway it is desired to assay.

Methods of recombinantly expressing chimeric G α proteins are known in the art and are described, for example, in Conklin et al., Nature 363:274-276 (1993), Komatsuzaki et al., FEBS Letters 406:165-170 (1995), and Saito et al., Nature 400:265-269 (1999). Additionally, chimeric G α proteins can be prepared by synthetic methods.

Signaling through G proteins can lead to increased or decreased production or liberation of second messengers, including, for example, arachidonic acid, acetylcholine, diacylglycerol, cGMP, cAMP, inositol phosphate and ions; altered cell membrane potential; GTP hydrolysis; influx or efflux of amino acids; increased or decreased phosphorylation of intracellular proteins; or activation of transcription. Thus, by using a chimeric G α subunit that binds U II receptor and couples to a desired signal transduction pathway in the methods of the invention, those skilled in the art can assay any convenient G protein mediated predetermined signal in response to U II and U II receptor agonists and antagonists.

Various assays, including high throughput automated screening assays, to identify alterations in G protein coupled signal transduction pathways are well known in the art. Various screening assay that measure

5 Ca⁺⁺, cAMP, voltage changes and gene expression are reviewed, for example, in Gonzalez et al., Curr. Opin. in Biotech. 9:624-631 (1998) and in Jayawickreme et al., Curr. Opin. Biotech. 8:629-634 (1997). Yeast cell-based bioassays for high-throughput screening of drug targets

10 for G protein coupled receptors are described, for example, in Pausch, Trends in Biotech. 15:487-494 (1997). A variety of cell-based expression systems, including bacterial, yeast, baculovirus/insect systems and mammalian cells, useful for detecting G protein coupled

15 receptor agonists and antagonists are described, for example, in Tate et al., Trends in Biotech. 14:426-430 (1996).

Assays to detect and measure G protein-coupled signal transduction can involve first contacting the

20 isolated cell or membrane with a detectable indicator. A detectable indicator can be any molecule that exhibits a detectable difference in a physical or chemical property in the presence of the substance being measured, such as a color change. Calcium indicators, pH indicators, and

25 metal ion indicators, and assays for using these indicators to detect and measure selected signal transduction pathways are described, for example, in Haugland, Molecular Probes Handbook of Fluorescent Probes and Research Chemicals, Sets 20-23 and 25 (1992-94). For

30 example, calcium indicators and their use are well known in the art, and include compounds like Fluo-3 AM, Fura-2, Indo-1, FURA RED, CALCIUM GREEN, CALCIUM ORANGE, CALCIUM CRIMSON, BTC, OREGON GREEN BAPTA, which are available from Molecular Probes, Inc., Eugene Oreg., and described,

for example, in U.S. Patent Nos. 5,453,517, 5,501,980 and 4,849,362.

Assays to determine changes in gene expression in response to a U II receptor agonist or antagonist can involve first transducing cells with a promoter-reporter nucleic acid construct such that a protein such as β -lactamase, luciferase, green fluorescent protein or β -galactosidase will be expressed in response to contacting U II receptor with U II. Such assays and reporter systems are well known in the art and are described, for example, at http://www.aurorabio.com/tech_platform-assay_technologies.html (visited August 5, 1999).

As described above, an assay to determine whether a candidate compound is a U II receptor agonist or antagonist is performed under conditions in which contacting the receptor with U II produces a predetermined signal. If desired, the assay can be performed in the presence of a given concentration of U II. Preferably, the U II concentration will be within 10-fold of the EC_{50} . Thus, an agonist that competes with U II for signaling through the U II receptor, or indirectly potentiates the signaling activity of U II, can be readily identified. Likewise, an antagonist that prevents U II from binding the U II receptor, or indirectly decreases the signaling activity of U II, can also be identified. Such compounds that demonstrate agonistic and antagonistic effects in the presence of U II are particularly useful for therapeutic applications, in which physiological concentrations of circulatory U II are likely to be present.

The invention also provides a method of identifying a U II receptor ligand. The method consists of contacting an isolated U II receptor with one or more candidate compounds in the presence of detectably labeled U II. A compound that reduces binding of the detectably labeled U II to the U II receptor is characterized as a U II receptor ligand. The invention also provides a composition that contains detectably labeled U II and an isolated U II receptor, which is useful, for example, in the method of identifying a U II receptor ligand.

As used herein, the term "detectably labeled U II" refers to U II derivatized with, or conjugated to, a moiety that is detectable by any analytical means. Detectably labeled U II useful in the methods disclosed herein generally retains its ability to bind U II receptor at subnanomolar concentrations. For example, a detectable moiety can be a radioisotope, fluorochrome, ferromagnetic substance, or luminescent substance. In one embodiment, the detectably labeled U II is radiolabeled. Exemplary radiolabels useful for labeling peptides include ^{125}I , ^{14}C and ^3H . Methods of detectably labeling peptides, either by incorporating labeled amino acids into the peptide during synthesis, or by derivatizing the peptide after synthesis, are known in the art. As described in Example III, below, an exemplary detectably labeled U II is human U II, radioiodinated at the core Tyr with ^{125}I , which binds membranes of cells transfected with U II receptor with an apparent K_d of 70 pM.

In order to determine whether a candidate compound decreases binding of detectably labeled U II to U II receptor, the amount of binding of a given amount of the detectably labeled U II is determined in the absence

of the candidate compound. Generally the amount of detectably labeled U II will be less than its K_d , for example, 1/10 of its K_d . Under the same conditions, the amount of binding of the detectably labeled U II in the presence of the candidate compound is determined. A decrease in binding due to a candidate compound characterized as a U II receptor ligand is evidenced by at least 2-fold less, preferably at least 10-fold to at least 100-fold less, such as at least 1000-fold less, binding of detectably labeled U II to U II receptor in the presence of the candidate compound than in the absence of the candidate compound.

An exemplary assay for determining binding of detectably labeled U II to U II receptor is the radioligand filter binding assay described in Example III, below. A variety of other low- and high-throughput assays suitable for detecting selective binding interactions between a receptor and a ligand are known in the art. Such assays include, for example, fluorescence correlation spectroscopy (FCS) and scintillation proximity assays (SPA) reviewed in Major, J. Receptor and Signal Transduction Res. 15:595-607 (1995); and in Sterrer et al., J. Receptor and Signal Transduction Res. 17:511-520 (1997)). Binding assays can be performed in any suitable assay format including, for example, whole cells or membranes that contain U II receptor, or substantially purified U II receptor polypeptide, either in solution or bound to a solid support.

A compound that is determined to be a U II receptor ligand in the methods of the invention can be assayed for its effect on U II receptor signaling, using the biological assays described herein. Thus, a "U II receptor ligand" can be determined to be a direct agonist

or antagonist of U II receptor, as described above, or can have little or no effect on U II receptor signaling in the absence of U II.

A U II receptor ligand that has little or no effect on U II receptor signaling is useful, for example, in many of the same applications as a U II receptor antibody. Thus, a U II receptor ligand can be used to specifically target a diagnostic or therapeutic moiety to a tissue that expresses U II receptor. In such applications, a U II receptor ligand can be labeled with a detectable moiety, such as a radiolabel, fluorochrome, ferromagnetic substance, or luminescent substance, and used to detect expression of U II receptor polypeptide in an isolated sample or in *in vivo* diagnostic imaging procedures. Likewise, a U II receptor ligand can be labeled with a therapeutic moiety, such as a cytotoxic or cytostatic agent or radioisotope, and administered in an effective amount to arrest proliferation or kill a cell or tissue that expresses U II receptor.

The U II receptor ligands, agonists and antagonists identified using the methods and compositions of the invention can be isolated and administered to an individual, such as a human or other mammal, in an effective amount to prevent or ameliorate a U II receptor-associated condition. As used herein, the term "U II receptor-associated condition" refers to any abnormal physiological or psychological condition in which a quantitative or qualitative alteration in signaling through the U II receptor contributes to the symptoms of the condition. A "U II receptor-associated condition" can also be any physiological or psychological condition in which altering signaling through the U II receptor has a beneficial effect in the individual.

Exemplary conditions for which administration of a U II receptor ligand, agonist or antagonist are contemplated include diseases of the circulatory system. As described in Example IV, the U II receptor is
5 expressed in cardiovascular tissues, and human U II acts as a vasoconstrictor in aorta. Thus, in instances where increased vasoconstriction is desired, administration of a compound that acts as a U II receptor agonist is beneficial. In contrast, in instances where vasodilation
10 is desired, administration of a compound that acts as a U II receptor antagonist is beneficial.

Accordingly, the U II receptor agonists, antagonists and ligands identified by the methods of the invention can be used to prevent or treat circulatory
15 disorders including, but not limited to, ischemia, hypertension, hypotension, angina pectoris, myocardial infarction, stroke, congestive heart failure, shock, stroke and impotence. Those skilled in the art understand which conditions are amenable to treatment
20 with agents having vasoconstrictive or vasodilatory activity.

Besides cardiovascular tissues, mRNA encoding the U II receptor has been observed in skeletal muscle, as well as in neural and sensory tissue, including
25 retina, circumvallate papillae (taste buds), olfactory epithelium, cerebellum, and choroid plexus (Tal et al., supra (1995)). Accordingly, uses of the agonists, antagonists and ligands identified by the methods of the invention in treating disorders of the nervous system,
30 muscle and eye are also provided. Disorders of the nervous system and muscle amenable to treatment by the compounds of the invention include disorders that result from trauma (e.g. spinal cord injury, hemorrhage),

infectious agents (e.g. meningitis, Scrapie, shingles),
immune malfunction (e.g. multiple sclerosis, AIDS,
myopathies, myasthenia gravis), neurodegeneration (e.g.
Parkinson's disease, Huntington's disease, amyotrophic
5 lateral sclerosis, Alzheimer's disease), poisoning (e.g.
carbon monoxide poisoning), inborn errors of metabolism
(e.g. leukodystrophies, neurometabolic storage
disorders), genetic disorders (e.g. muscular dystrophies)
and tumors (e.g. glioblastomas and astrocytomas), as well
10 as acute and chronic pain. Disorders of the eye,
amenable to treatment by the methods of the invention
particularly include diseases of the retina, such as
diabetic retinal disease and retinoblastoma.

Additionally, the agonists, antagonists and
15 ligands identified by the methods of the invention are
useful in treating psychiatric disorders, such as
depression; anxiety disorders, such as generalized
anxiety disorder, panic attacks, obsessive-compulsive
disorder, phobias, acute stress disorder, post-traumatic
20 stress disorder; and psychotic disorders, such as
unipolar mania or depression, bipolar disorder and
schizophrenia. Furthermore, the agonists, antagonists
and ligands identified by the methods of the invention
are useful in treating disorders of behavior, memory and
25 learning, and sleep. Disorders of behavior include, but
are not limited to, autistic disorder, Asperger's
disorder, aggression, pervasive developmental disorders,
Tourette's syndrome, attention-deficit hyperactivity
disorder and addiction. Disorders of memory and learning
30 include, but are not limited to, dementia, including
dementia due to neurodegenerative diseases, infectious
disease, proliferative diseases, endocrine disease,
tumors, metabolic disorders, and toxins; and
developmental learning disabilities. Disorders of sleep

and of the sleep-wake cycle include, but are not limited to, insomnia, bedwetting, sleepwalking, sleep apnea and narcolepsy.

The U II receptor agonists, antagonists and
5 ligands of the present invention can be conveniently formulated for therapeutic administration together with a pharmaceutically acceptable carrier. Suitable pharmaceutical carriers for the methods of the invention are well known and include, for example, aqueous
10 solutions such as physiologically buffered saline, and other solvents or vehicles such as glycols, glycerol, oils or injectable organic esters. A pharmaceutical carrier can contain a physiologically acceptable compound that acts, for example, to stabilize or increase the
15 solubility of a pharmaceutical composition. Such a physiologically acceptable compound can be, for example, a carbohydrate, such as glucose, sucrose or dextrans; an antioxidant, such as ascorbic acid or glutathione; a chelating agent; a low molecular weight protein; or
20 another stabilizer or excipient. Pharmaceutical carriers, including stabilizers and preservatives, are described, for example, in Martin, Remington's Pharm. Sci., 15th Ed. (Mack Publ. Co., Easton, 1975).

Those skilled in the art can formulate the
25 therapeutic compounds to ensure proper distribution in vivo. For example, the blood-brain barrier (BBB) excludes many highly hydrophilic compounds. To ensure that the therapeutic compounds of the invention cross the BBB, they can be formulated, for example, in liposomes,
30 or chemically derivatized. Those skilled in the art understand that the choice of the pharmaceutical formulation and the appropriate preparation of the

composition will depend on the intended use and mode of administration.

Methods of introduction of a therapeutic compound of the invention include, but are not limited to, intradermal, intramuscular, intraperitoneal, intravenous, subcutaneous, oral, intranasal, intraspinal and intracerebral routes. Methods of introduction can also be provided by rechargeable or biodegradable devices, particularly where gradients of concentrations of drug in a tissue is desired. Various slow release polymeric devices are known in the art for the controlled delivery of drugs, and include both biodegradable and non-degradable polymers and hydrogels.

An effective dose of a therapeutic compound of the invention can be determined, for example, by extrapolation from the concentration required for modulating U II receptor signaling or binding in *in vitro* assays described herein, or from the dose required for efficacy in the aorta vasoconstriction assay described herein. Typically, an appropriate dose of a therapeutic compound of the invention can be in the range of 0.001-100 mg/kg of body weight, such as in the range of 0.01-10 mg/kg of body weight. An appropriate dose can be determined by those skilled in the art based on the bioactivity of the particular compound, the desired route of administration, the gender and health of the individual, the number of doses and duration of treatment, and the particular condition being treated.

The following examples are intended to illustrate but not limit the present invention.

EXAMPLE I**Identification of Urotension II as the
endogenous ligand of GPR14**

This example shows the identification of
5 Urotensin II (UII) as the endogenous ligand for the
orphan G-protein coupled receptor GPR14. This example
also shows a signaling assay that can be used in a method
of identifying a U II receptor agonist or antagonist.

In order to identify the natural ligand of
10 GPR14, peptide extracts from a variety of different
mammalian tissues were prepared. The peptide extracts
were prepared by mincing, under liquid nitrogen, tissue
obtained from a slaughterhouse and immediately boiling
the tissue for 10 min. in deionized water at a 1:2
15 tissue/water (g/ml) ratio. After cooling to 10°C, acetic
acid was added to 1M final concentration. The homogenate
was further treated with a Polytron for 2 min. After
centrifugation for 30 min at 12,000 x g, the supernatant
was removed, and the pellet re-extracted with one volume
20 of 1M acetic acid. The supernatants were combined, and
three volumes of acetone were added. The precipitate was
removed at 15,000 x g for 30 min. After concentration of
the supernatant by rotor evaporation, the remainder was
extracted twice with two volumes of ethyl ether, then
25 frozen and lyophilized. The lyophilized material was
resuspended in 5% CH₃CN, 0.1% trifluoroacetic acid (TFA)
and applied on a PrepPak-Delta-Pak C18, 15 µm, 300 Å,
25 x 100 mm (Waters) preequilibrated with 5% CH₃CN/0.1%
TFA. The material was eluted with a linear gradient from
30 5% to 39% CH₃CN/0.1% TFA, with the active fractions
eluting at 34%-38% CH₃CN.

The tissue extracts were fractionated by preparative reverse-phase high performance liquid chromatography (rpHPLC) into 72 individual one-minute fractions. Aliquots of each fraction were tested for induction of intracellular calcium changes using Chinese hamster ovary cells transiently transfected with GPR14 cDNA and loaded with the calcium-sensitive dye fluo-3. Changes in fluorescence were monitored over 240 seconds using a fluorescence imaging plate reader (FLIPR) system, following procedures described by Coward et al., Proc. Natl. Acad. Sci. USA 95:352-357 (1998). The maximal fluorescence increment generated by each fraction on GPR14 transfected cells was normalized to the maximal control value seen in mock-transfected cells.

As shown in Figure 2A, a robust calcium change was observed in two adjacent rpHPLC one-minute fractions of a bovine hypothalamic tissue extract, fractions 53 and 54. A similar change could not be detected either in non-transfected cell lines or in cells transfected with other orphan receptors.

The highest levels of activity were detected in bovine hypothalamic tissue, which was consequently used for large-scale purification. The active component was purified over a seven-step purification strategy combining reverse-phase and cationic-exchange HPLC columns. One single activity peak was detected, indicating that the activity can be attributed to a unique molecular entity. The bioactive compound was found to be extremely scarce, precluding a total structural analysis. However, the calcium response to the active material depicted a distinctive time course, as shown by the open circles in Figure 2B.

The active material was determined to be sensitive to trypsin and reducing agents, indicating that the biological activity could be attributed to a peptide containing basic amino acid(s) and cysteine residues.

- 5 Because GPR14 is distantly related to the somatostatin receptors (SSTRs), sharing the highest similarity to SST4R (41% in the transmembrane regions), somatostatin-like, cysteine bridge-containing peptides were screened in the above system for their ability to activate calcium
- 10 signaling through GPR14. The peptides tested were human melanin-concentrating hormone (MCH) (SEQ ID NO:7), somatostatin-14 (SST-14) (SEQ ID NO:8), cortistatin-14 (SEQ ID NO:9) and goby urotensin II (U II) (SEQ ID NO:4) (see Table 1). While MCH was inactive, SST-14 and the
- 15 related peptide cortistatin were shown to activate GPR14, albeit only at very high concentrations (see Table 1). In contrast, UII induced a robust increase in $[Ca^{2+}]_i$ with a median effective concentration (EC_{50}) of 0.14 nM (see Table 1), demonstrating potency more than 25,000-fold
- 20 greater than SST-14.

U II was originally isolated from goby, as described by Pearson et al., Proc. Natl. Acad. Sci. USA 77:5021-5024 (1980), and an expressed sequence tag has been described by Coulouarn et al., Proc. Natl. Acad. Sci. USA 95:15803-15808 (1998) coding for a putative mammalian counterpart expressed in humans. The human peptide containing a single cysteine bridge (SEQ ID NO:5; Table 1), as proposed from the structure of the human -urotensin II precursor, was synthesized and tested on GPR14 expressing CHO cells.

The time course of the calcium response caused by the synthetic peptide at 0.1 nM (solid circles in Figure 2B) was identical to that observed with an active fraction isolated from bovine hypothalamic tissue extracts (open circles in Figure 2B). This result indicated that both synthetic mammalian U II, as well as the natural compound fractionated from bovine hypothalamic tissue, act on the same receptor.

It was also determined that an active compound from brain extracts of squirrel monkeys eluted by rpHPLC with the same retention time as synthetic human U II peptide, further indicating that the monitored activity is caused by U II. The activity found in bovine tissue eluted slightly earlier than human U II, suggesting some structural differences among the mammalian U II peptides.

In summary, these results indicate that GPR14 is the Urotensin II receptor. These results also demonstrate that the signaling activity of U II can be determined by monitoring calcium influx in cells transfected with GPR14.

EXAMPLE II**Pharmacological characterization of Urotension II**

This example shows the specificity and potency of U II for the Urotensin II receptor, GPR14. This example also shows a binding assay that can be used in a method of identifying a U II receptor ligand.

In order to further define the pharmacological profile of GPR14, peptides structurally related to U II were tested for their ability to activate GPR14 expressing cells. Goby U II, which is identical to human U II at the C-terminus but differs in its N-terminus, was equipotent to human U II, as shown in Figure 3A and Table 1), pointing at the cyclic ring-structure as an important binding motif. Based on molecular modeling data, a putative U II antagonist [D-Trp⁷]U II-(5-10) (SEQ ID NO:6, see Table 1) has been proposed (Perkins et al. Biochem. Soc. Trans. 18:918-919 (1990))). In the GPR14 transfected CHO cell system, [D-Trp⁷]U II-(5-10) behaved as an agonist with a 190-fold lower potency compared to human U II, but with no antagonistic activity at concentrations up to 1 μ M.

Somatostatin-14 and the somatostatin analogs RC-160 (SEQ ID NO:10) and octreotide (SEQ ID NO:11) were also tested for their ability to activate GPR14 (see Table 1). RC-160 activated the receptor with an EC₅₀ of 338 +/- 53.6 nM. Octreotide, structurally similar to RC-160, showed even lower potency. Somatostatin-14 was active on GPR14 expressing cells but at physiologically non-relevant concentrations (see Figure 3A and Table 1). This activity was not observed in untransfected cells. Cortistatin-14, which differs from somatostatin-14 in a C-terminal lysine residue, exhibited even lower affinity,

pointing out the importance of having a non-charged carboxy terminus for GPR14 reactivity.

In contrast, as shown in Table 1, human U II was unable to activate the somatostatin 2A receptor (SSTR2A) transiently expressed in HEK 293T cells in the presence of the chimeric GTP-binding protein Gαq/i3, whereas RC-160 and octreotide mobilized $[Ca^{2+}]_i$ in these cells with a high affinity similar to published values (Patel et al., Life Sci. 57:1249-1265 (1995)).

In summary, these results demonstrate that U II activates the Urotensin II receptor, GPR14, with high potency and high specificity.

EXAMPLE III

Binding of Urotension II to GPR14

This example shows that human U II binds to GPR14 with high affinity and in a saturable manner.

In order to develop a U II analog that would be suitable for binding studies, human U II was initially labeled with cold iodine. Human U II contains one tyrosine residue at position 9 that can be used for iodination. Tyr⁹ was iodinated with NaI according to the method of Markwell et al., Anal. Biochem. 125:427-432 (1982), using IODO-BEADS® (Pierce, Rockford, IL).

Monoiodo-hU II was shown to induce a concentration-dependent transient increase in $[Ca^{2+}]_i$ in CHO cells stably transfected with GPR14 with an EC₅₀ similar to the non-modified peptide (Figure 3A, Table 1). In contrast, the diiodo-form of hU II showed a 25-fold decrease in potency compared to hU II. Consequently, the

monoiodinated form of human U II was used as a radioligand for the determination of physical constants. [^{125}I -Tyr⁹]hU II was prepared as described above, using 0.5 nmol hU II and 500 μCi Na ^{125}I (Amersham).

5 Ligand binding assays were performed in 96-well glass fiber type B filtration plates using a MultiScreen filtration system (Millipore) pre-coated with 0.5% BSA at room temperature for one hour. CHO cells stably
10 expressing GPR14 were harvested with a cell scraper and resuspended in binding buffer containing 20 mM Tris-HCl, pH 7.4, 2 mM MgCl₂, 0.25% BSA and 0.25 mg/ml bacitracin. Membranes were prepared using a Polytron tissue homogenizer. Total membrane particulate was obtained after centrifugation at 20,000 x g for 60 min at 4°C. The
15 resulting pellet was passed through a 27-gauge needle five times and the membranes were stored at -80°C until use. Protein concentration was determined using a BCA Protein Assay Kit (Pierce). Membranes containing 5 μg total protein were incubated with monoiodinated
20 [^{125}I -Tyr⁹]hU II for 60 min in a total volume of 200 μl of binding buffer. Filters were washed with cold phosphate buffered saline three times and radioactivity was counted using a MicroBeta liquid scintillation counter (EG&G Wallac, Gaithersburg, MD). Non-specific binding was
25 determined in the presence of 1 μM unlabeled human U II. Data were analyzed using PRISM™ software (GraphPad, San Diego, CA).

Figure 3B shows saturation binding of ^{125}I -labeled hU II. Membranes of CHO cells stably
30 expressing GPR14 were incubated with various concentrations of ^{125}I -labeled hU II and the bound ligand separated by filtration, as described above. Concentrations of free ligand were calculated by

subtracting the amount of specifically bound ligand from the total amount of radioligand added. The data represent the average of two independent experiments done in triplicate. The insert shows a Scatchard transformation of the specific binding data. As shown in Figure 3B, [¹²⁵I-Tyr⁹]hU II displayed saturable and displaceable binding to membranes of cells transfected with GPR14 with an apparent dissociation constant (K_d) of 70 pM and a maximal binding capacity (B_{max}) of 350 fmol per milligram of membrane protein.

EXAMPLE IV

Expression profile and activation of GPR14 in cardiovascular tissue

This example shows the distribution of GPR14 mRNA in cardiovascular tissue and the physiological activation of GPR14 by human U II.

As demonstrated above, activation of GPR14 by U II stimulates Ca²⁺ influx, suggesting a possible physiological role for GPR14 as an excitatory receptor, in contrast to somatostatin receptors, which are coupled to inhibition of adenylate cyclase. Physiological actions of fish U II have been described in the mammalian cardiovascular system (Gibson et al., Gen. Comp. Endocrinol. 64:435-439 (1986)), and specific fish U II binding sites on rat thoracic aortic membranes have been reported (Itoh et al., Eur. J. Pharm. 149:61-66 (1988)). The expression of GPR14 mRNA in various tissues of the cardiovascular system was tested using a qualitative reverse transcription-polymerase chain reaction (RT-PCR) approach. Reverse transcription reactions were performed with Superscript II reverse transcriptase (Life Technologies) using 5 µg of total tissue RNA. 10% of the

final reaction products were used in PCR reactions containing 0.2 μ M each of primers having the sequences 5'-CTGAGCCTGGAGTCTACAACAAGC-3' (SEQ ID NO:12) and 5'-TAGGTGGCTATGATGAAGGGAATG-3' (SEQ ID NO:13). Reactions
5 were carried out using Goldtaq polymerase (Perkin Elmer) and buffer conditions recommended by the manufacturer. The PCR conditions were 94°C for 10 min, followed by 34 cycles of 94°C for 30 sec, 58°C for 30 sec and 72°C for 60 sec. Reaction products were separated on a 1% agarose
10 gel.

As shown in Figure 4A, GPR14 is expressed in rat thoracic aorta (lane 2), abdominal aorta (lane 3), and heart (lane 4). Because GPR14 was shown to be expressed in thoracic aorta, hU II was tested for its
15 ability to cause vascular smooth muscle contractions in thoracic aorta.

To prepare thoracic aorta tissue segments, four month old male Fischer 344 rats were sacrificed by decapitation and their thoracic aorta removed. Arteries
20 were cut into 3 mm segments and mounted on platinum wires in an oxygenated tissue bath containing 40 ml of Krebs' solution (118 mM NaCl, 5.2 mM KCl, 1.6 mM CaCl_2 , 1.2 mM KH_2PO_4 , 25.5 mM NaHCO_3 , 1.2 mM MgSO_4 , 0.027 mM disodium EDTA, and 115 mM glucose), and kept at 37°C and oxygenated
25 with 95% O_2 /5% CO_2 . Tissue segments were slowly stretched to 1 gram resting tension, then allowed to equilibrate for 1 hour, with fresh Krebs' solution added every 20 min. Contractile responses to cumulative addition of urotensin (10^{-10} to 10^{-7} M) were recorded using Fort 10
30 force transducers and a MacLab analog-digital converter. Maximal contraction of each arterial segment was then determined by addition of 10^{-4} M norepinephrine, and endothelial integrity confirmed by measuring relaxation

to acetylcholine (10^{-6} M). The contraction induced by human U II is expressed in Figure 4B as the percentage of the maximum change in tension induced by norepinephrine (NE), with the bars indicating the S.E.M of four independent experiments.

As shown in Figure 4B, human U II was found to act as a vasoconstrictor in a concentration dependent manner with an EC_{50} of 2.4 nM. Maximal activity reached 32% of norepinephrine-induced contraction. The response was characterized by a slow onset (10-15 min) and a long lasting maintained tone, which could not be washed out even after one hour of frequent changes of the bath solution.

Contraction of aortic smooth muscle cells by fish U II has been shown to be calcium dependent and to be blocked by the calcium channel blocker nitrendipine. Furthermore, as shown in Example I, GPR14 activation by U II induces a strong Ca^{2+} influx in transfected cells. Therefore, it is concluded that GPR14 is the receptor responsible for hU II-directed contraction of rat thoracic aorta.

EXAMPLE V

Expression profile and processing of human preprourotensin II

This example shows the expression of human preprourotensin II in human kidney, and the ability of kidney cells to process preprourotensin II into biologically active urotensin II.

In order to determine where circulating human U II is produced, a normalized Northern dot blot was

probed with full length human preprorenin II
(pphU II).

The tissues shown on the human masterblot
(Clontech) in Figure 5A are: A1, whole brain; A2,
5 amygdala; A3, caudate nucleus; A4, cerebellum; A5,
cerebral cortex; A6, frontal lobe; A7, hippocampus; A8,
medulla oblongata; B1, occipital lobe; B2, putamen; B3,
substantia nigra B4, temporal lobe; B5, temporal lobe;
B6, subthalamic nucleus; B7, spinal cord; C1, heart; C2,
10 aorta; C3, skeletal muscle; C4, colon; C5, bladder; C6,
uterus; C7, prostate; C8, stomach; D1, testis; D2, ovary;
D3, pancreas, D4, pituitary; D5, adrenal; D6, thyroid;
D7, salivary; D8, mammary gland; E1 kidney; E2, liver;
E3, small intestine; E4, spleen; E5, thymus; E6,
15 peripheral leukocytes; E7, lymph nodes; E8, bone marrow;
F1, appendix; F2, lung; F3, trachea; F4 placenta; G1,
fetal brain; G2, fetal heart; G3, fetal kidney; G4, fetal
liver; G5, fetal spleen; G6 fetal thymus; G7 fetal lung;
H1 yeast total RNA, H2, yeast tRNA; H3, *E. coli*, tRNA;
20 H4, *E. coli* DNA; H5, poly r(A) RNA, H6, human C₆t1 DNA; H7
human DNA (100ng); H8 human DNA (500 ng). Hybridization
conditions were as recommended by the manufacturer.

Of the 50 different human tissues, pphU II mRNA
was found to be expressed at high levels only in human
25 kidney (see Figure 5A). Moderate to low expression was
found in spinal cord and medulla oblongata respectively,
but these are not expected to be a source of circulating
U II. All other tissues examined either expressed none
or very low amounts of the UII precursor. The expression
30 and processing of the UII precursor potentially varies
from individual to individual, which may be a reflection
of the overall health and condition of the individual.

Based on the high expression of pphU II mRNA in kidney, an investigation was made of whether biologically active hU II could be secreted by cells derived from the kidney. The human U II precursor cDNA was obtained from
5 Research Genetics (IMAGE clone # 926809) and sequenced from both ends. The insert which contained the full open reading frame was subcloned into pcDNA3.1(+) (Invitrogen, San Diego). The expression construct was transfected into human embryonic kidney cells (HEK 293T cells) grown
10 in DMEM containing 10% fetal bovine serum. For peptide expression studies the medium was replaced by Opti-MEM (Life Technologies) with or without added serum, and the cells grown for 24 hours.

The conditioned medium from HEK 293T cells
15 transfected with pphU II cDNA was removed from the cell layer and applied to a SePak C18 cartridge (Waters). Absorbed material was eluted with 60% CH₃CN/0.1% TFA and fractionated using a Nova-Pak C18 3.9x150 mm analytical column (Waters) equilibrated with 5% CH₃CN/0.1% TFA. The
20 material was eluted with a linear gradient to 60% CH₃CN/0.1% TFA over 60 min. One minute fractions were collected and tested on CHO cells stably expressing GPR14 for their ability to induce Ca²⁺ influx as described above. Non-transfected HEK 293T cells were treated under
25 identical conditions and used as a negative control.

As shown in Figure 5B, activity was detected in fractions which elute at the same position as synthetic hU II, indicating that the active peptide is expressed, proteolytically processed and properly cyclized in this
30 *in vitro* system. Conditioned medium of mock-transfected HEK cells treated under identical conditions failed to produce any response in GPR14 transfected cells. The presence of serum in the medium increased the quantity of

active material produced by the HEK 293T cells, suggesting a possible serum component actively involved in processing of the precursor.

5 These results demonstrate that kidney may be the peripheral source of human U II which modulates vascular function. These results also demonstrate that human U II can be processed from its precursor in the same molecular form as that the biologically active peptide shown as SEQ ID NO:5 in Table 1.

10 All journal article, reference and patent citations provided above, in parentheses or otherwise, whether previously stated or not, are incorporated herein by reference in their entirety.

15 Although the invention has been described with reference to the examples provided above, it should be understood that various modifications can be made without departing from the spirit of the invention. Accordingly, the invention is limited only by the claims.

We claim:

1. A method of identifying a urotensin II (U II) receptor agonist or antagonist, comprising:
 - (a) contacting an isolated U II receptor with
5 one or more candidate compounds under conditions wherein said U II receptor produces a predetermined signal in response to U II; and
 - (b) identifying a candidate compound that
alters production of said signal, said compound being
10 characterized as a U II receptor agonist or antagonist.
2. The method of claim 1, wherein said predetermined signal is calcium ion influx.
3. The method of claim 1, wherein said one or
more candidate compounds comprises greater than 10^5
15 compounds.
4. The method of claim 1, wherein said U II
receptor comprises the amino acid sequence of SEQ ID
NO:2, or modification or fragment thereof having U II
receptor activity.
- 20 5. The method of claim 1, wherein said U II
is human U II.
6. The method of claim 1, wherein said
candidate compound contacts U II receptor in the presence
of U II.

7. A method of identifying a U II receptor ligand, comprising:

(a) contacting an isolated U II receptor with one or more candidate compounds in the presence of
5 detectably labeled U II; and

(b) identifying a compound that decreases binding of said detectably labeled U II to the U II receptor, said compound being characterized as a U II receptor ligand.

10 8. The method of claim 7, wherein said one or more candidate compounds comprises greater than 10^5 compounds.

9. The method of claim 7, wherein said U II receptor comprises the amino acid sequence of SEQ ID
15 NO:2, or modification or fragment thereof having U II receptor activity.

10. The method of claim 7, wherein said detectably labeled U II is human U II.

11. The method of claim 7, wherein said
20 detectably labeled U II is radiolabeled U II.

12. The method of claim 7, wherein said radiolabel is ^{125}I .

13. A composition, comprising an isolated U II receptor and detectably labeled U II.

25 14. The composition of claim 13, wherein said wherein said U II receptor comprises the amino acid sequence of SEQ ID NO:2, or modification or fragment thereof having U II receptor activity.

15. The composition of claim 13, wherein said detectably labeled U II is human U II.

16. The composition of claim 13, wherein said detectably labeled U II is radiolabeled U II.

5 17. The composition of claim 16, wherein said radiolabel is ^{125}I .

```

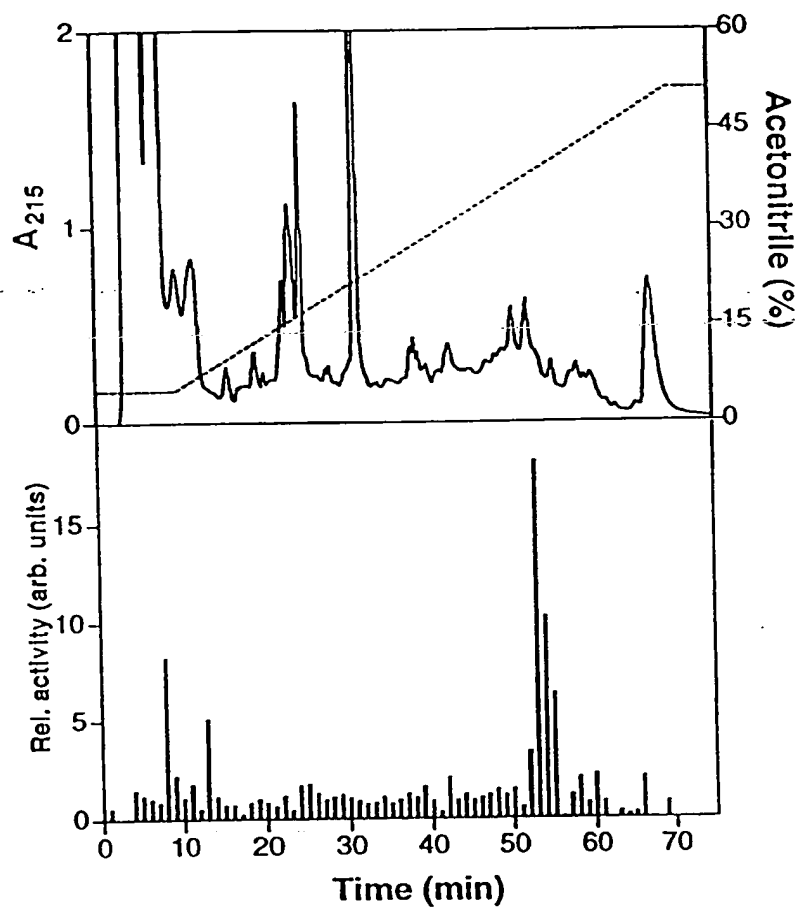
1 atgcttttga tgcactgccg tccgtcactg agtatgtaat gtgaggtaga gatccaggcc
61 catttccttc atataacaaa tggaaaccca aaggcctttt atctattttt ctatttgaac
121 ttttacaccc agcctagctg ttacttttgt gattggaaga aatctttttt aaaaaattaa
181 tttattttatt ttatatatgt gagcacactg tcatcactgt cttcatcact gtcttcagac
241 acaccagaaa aatgcattgg acctcttaca gatggttggt agccaccatg tgggtgctgg
301 gaattgaact caggacctct ggaagagcag gcagtgttct taaccgcgtg acatctttcc
361 agaccaagaa atcatttttt aaaaacacac acgtaacgta aaatgggaaa tgggactctc
421 tctgaggctg tggagctaga acttaggact ctttgccacac tctgacctct gcctcactgt
481 ggccaacggt tcagaaaaag catttggttg agacttgggg ttctactggg gagcgtcctg
541 gggacagtgg gtcccaatgg ctctagggtc ctctgtgta gctggggaga taacacaaaa
601 gggattcttt tgaggcttcc aacaggagat aggacctgtt gagcctttgt ctctctgcat
661 agggacagtg actgtgtcca tcacagaggc tgtttagggtc atagaagtag gttactgect
721 tgaacctctg acactaatct tttcccacag gacaagtttc ccacgggctc tctcactga
781 gcagtgggtc tccccctgga atcccagtgt gaggaccgag atggctctga gcctggagt
841 tacaacaagc tttcatatgc tcaccgtgtc cggaaagcact gtgactgagc tgcctggtga
901 ctccaacgtg tccctcaaca gtctctgggt cggcccaaca gatccagct cctgaaaga
961 ccttggtggc acgggtgtca tccgggcagt gctctcagcc atgggtgtgg tgggcatggt
1021 gggaaatgta tacactttgg tggcatgtg cgggtttctg cgtgcctcgg cctccatgta
1081 cgtctatgtg gtcaacctag cgctggctga tctgctgtac ctgctgagca ttccttcat
1141 catagccacc tacgtcacta aggactggca ctttgagatg gtgggctgca gactcctctt
1201 tagcctggac ttcttgacaa tgcacgccag catcttcacc ctgaccataa tgagcagcga
1261 acgctatgca gccgtactga ggcctctgga cacagtccag cgctccaagg gttaccgtaa
1321 gctgctggtg ctgggcacct ggttgctggc actgctgctg accctaccca tgatgcttgc
1381 catccagctg gtccgcaggg gctctaagag cctctgectg ccagcctggg gccctcgtgc
1441 ccaccgtact tacctaacgt tgctctttgg gaccagcatt gtggggcctg gcttggctcat
1501 tgggctgctc tatgtccgtc tggccagggc gctctacctc atccttggtg tgcctcttct
1561 gcagacacgg cggtctgccc accccagggt gctctacctc ctggccaggt accacgaggc
1621 cttctggggc tgctttctac cttcttggtc gtggcagctg ctggccaggt accacgaggc
1681 catgccactg actcccgaga ctgcacgcat tgtcaactac ctgaccacct gcctcactta
1741 tggcaacagt tgcatcaatc ctttctctca cactctgctc accaagaact atcgagtaga
1801 cctacgtggc cgccagcggg cactgggtag tagttgccac agcccaggga gtccctggcag
1861 ctctctgccc agccgagctc acctccagca ggactcgggc cgctcgctgt cctccagcag
1921 ccaacaggcc acagagaccc tcatgctgtc tccagtcacc cgtaacgggg cccttctctg
1981 agagtgcact gtgcaatcct ggcataaggaa aggaccctaa ggcgtgctgc tccggagcgc
2041 attcccagaa tccccctgctc aaacctaaat ggctcgtctg tcttcttcca gttctcttct
2101 ggaaagcccc ctggtgtttc ttctgtgctc acttcagctc tttcgccaac tttgttacga
2161 tgccaacaca tacacaacca atctctgagg gacttcggaa gctgtggtgt gccgagtcac
2221 ctaaggccta attgcccaga gcagctagta caggctgacc tgtgatgtaa gcccttccag
2281 ccctgaggac cttcccggac tgggtggtaag ggaccggcag ggtatctcta cctctctccc
2341 atcccatggt agcttcttgc ttccgtatgc acatggaaca ataaaccatg tactctgcag
2401 tgtgtcctgc ctcaagtgtg gggatggaag ttgtggtggg gtggaaatcc tgaactctgc
2461 gatccccagc atgaggaacc taagtgtttg gtgaggttcc caaacctttg cctgtggatg
2521 tccctacatg gttgtccttg ggacagcttt tcttggtgct actgtttctc ttaccacggg
2581 taaaagacag ctgcccaccc tcttcagggt gaccatcaga ggggtgggacg gcacctgggt
2641 ctggttagtac tcttgctgaa tgtggcaggg catggagtgc aggtctctgc agcgacaaag
2701 tgttccagaa gtagattctg gtgtctgacc acacacgcag agagagacac acgcagagag
2761 aaaatgagag ataccatac ccacatgcgc acacacgcag agagagacac acgcagagag
2821 agacacacgc agagagaaat gaaggggttt agcctgtaat tccaggagtt ggggaggctt
2881 gaggtcagcc tgtgctactt attaaagact ttaaatttaa aaaaaaaac aggggtgggg
2941 atttagctca gtggtagagc gcttgccctg aaagcgcaag gccctgggtt cggctcctcag
3001 ctccgaaaaa aagaaaaaga aaaacaaaca aacaaacttt tctttgagac aggtctctcag
3061 tatgcagcct tggctggcct tagacttact gtgtagacaa ggctggcctg atctcacaga
3121 ga

```

MALSLESTTSFHMLTVSGSTVTELPGDSNVSLNSSWSGPTDPSS
 LKDLVATGVIGAVLSAMGVVGMVGNVYTLVVMCRFLRASASMYVYVNLALADLLYLL
 SIPFYIATYVTKDWHFGDVGCRLVFLSLDFLTMHASIFTLTIMSSERYAAVLRPLDTVQ
 RSKGYIKLLVLGTLWLLALLLTPMMLAIQLVRRGSKSLCLPAWGPRAHRTYLLFLGT
 SIVGPGVLVIGLLYVRLARAYWLSQOASFQTRRLPNPRVLYLILGIVLLFWACFLPFW
 LNLQALQYHEAMPLTPETARIVNYLTTCLTYGNSCINPFLYTLTKNYREYLRGRQRS
 LGSSCHSPGSPGSFLPSRVHLQQDSGRSLSSSSQATETLMLSPVPRNGALL

FIGURE 1

A



B

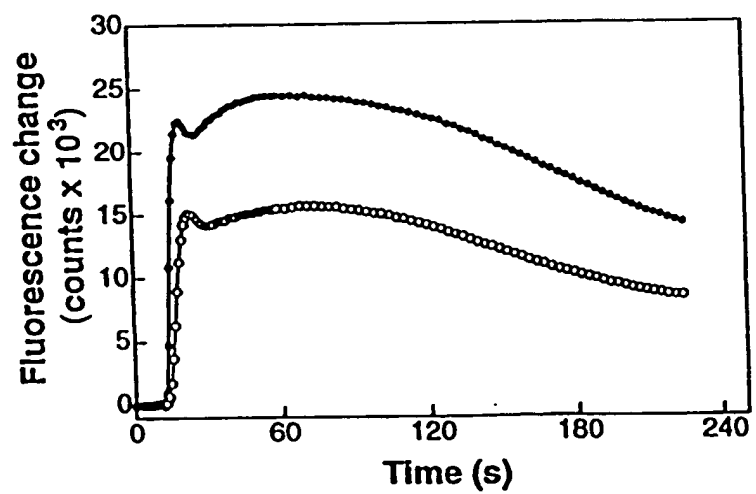


FIGURE 2

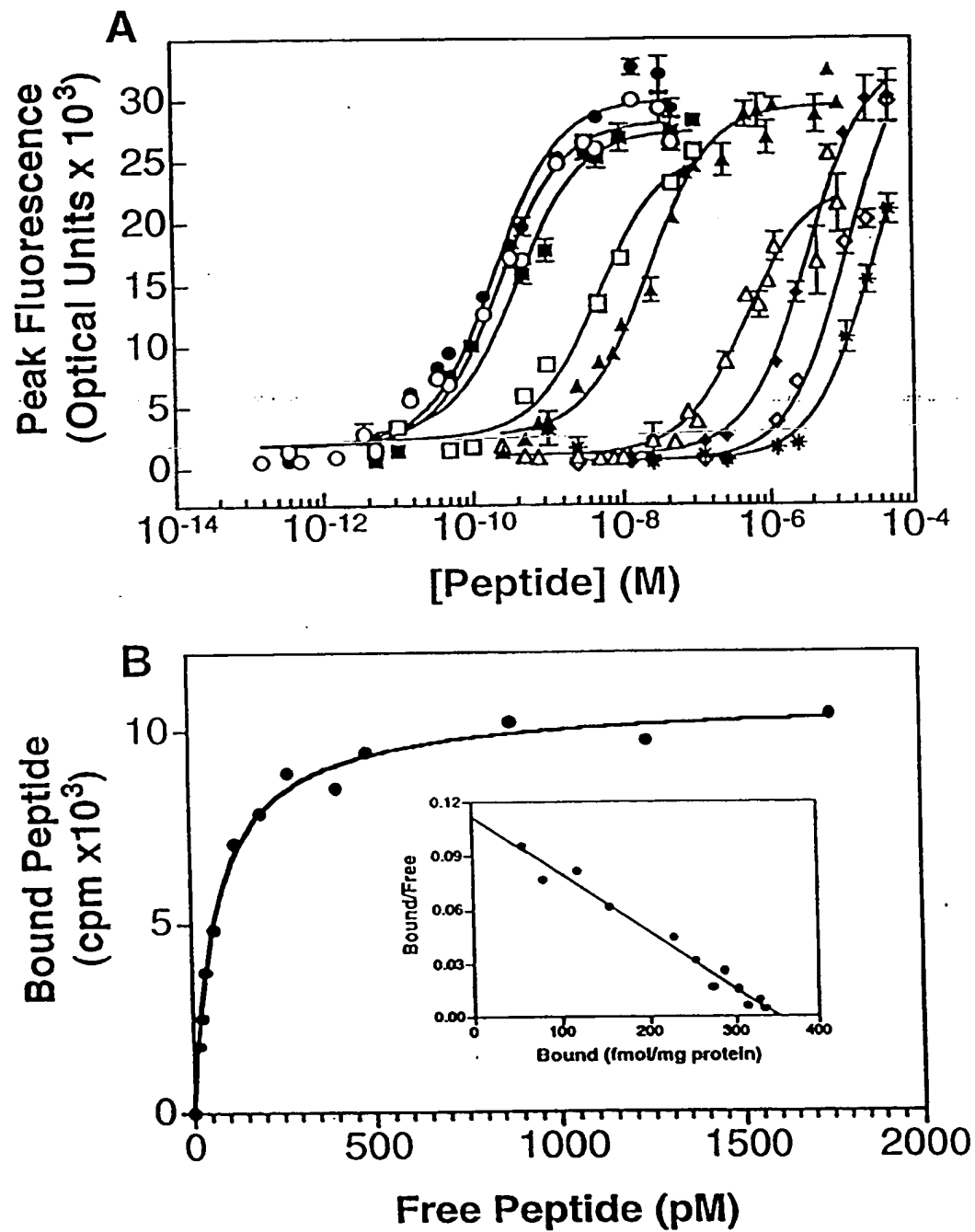


FIGURE 3

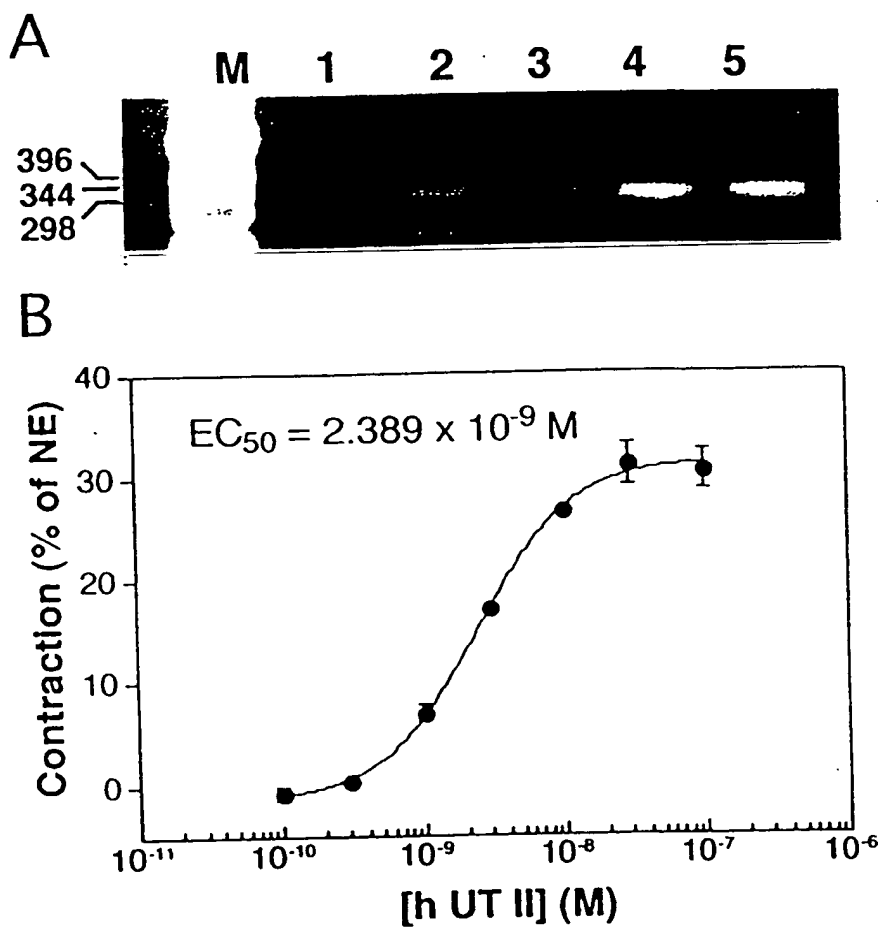


FIGURE 4

5/5

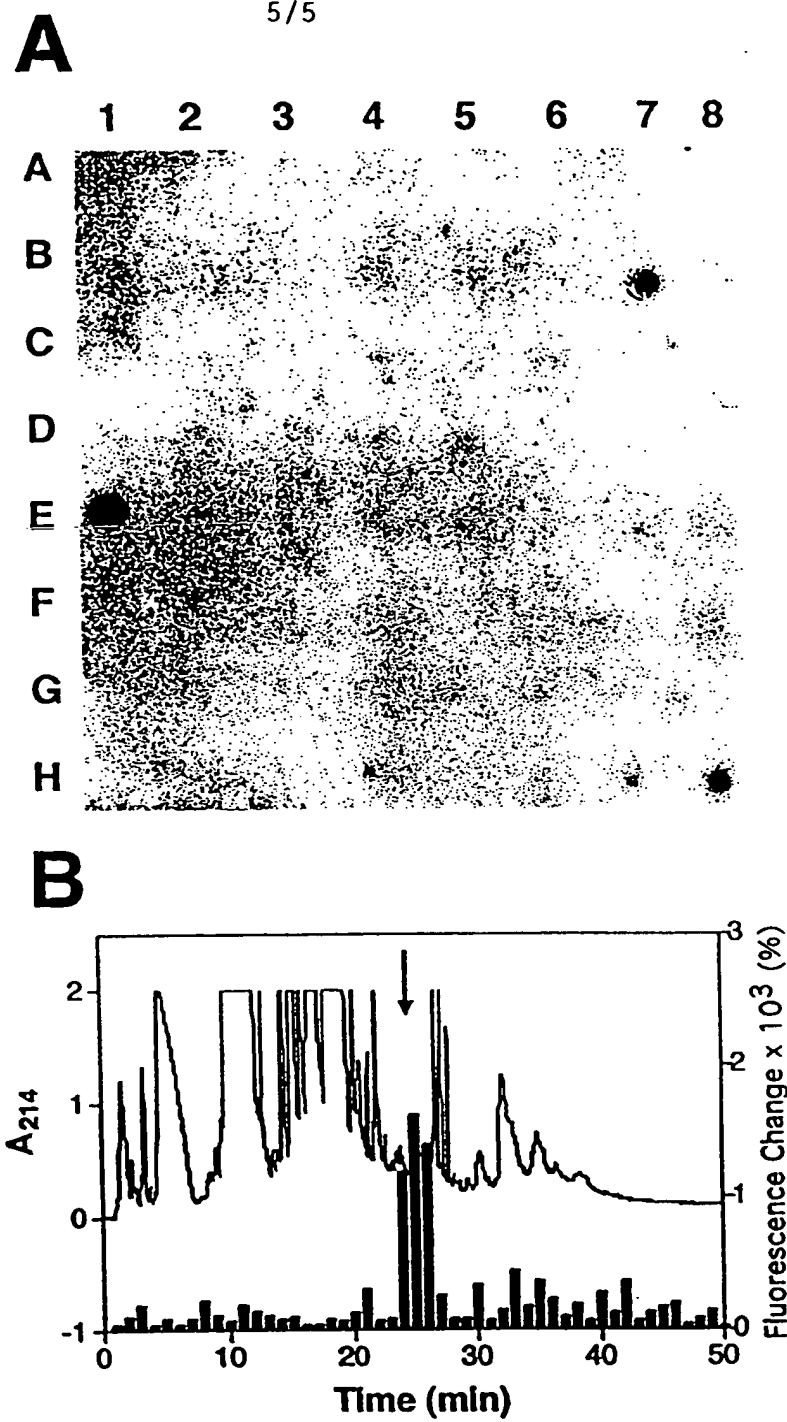


FIGURE 5

SEQUENCE LISTING

<110> Civelli, Olivier
Nothacker, Hans-Peter

<120> Urotensin II Receptor

<130> P-UC 3535

<140> 09/379,968

<141> 1999-08-24

<160> 12

<170> PatentIn Ver. 2.0

<210> 1

<211> 3122

<212> DNA

<213> Rat

<220>

<221> CDS

<222> (821)..(1981)

<400> 1

atgcttttga tcgactgccg tcggtcactg agtatgtaat gtgaggtaga gatccaggcc 60
catttccttc atataacaaa tggaaaccca aaggcctttt atctattttt ctatttgaac 120
ttttacaccc agcctagctg ttacttttgt gattggaaga aatctttttt aaaaaattaa 180
tttattttatt ttatatatgt gaggacactg tcatcactgt cttcatcact gtcttcagac 240
acaccagaaa aatgcattgg acctcttaca gatgggttggt agccaccatg tggttgctgg 300
gaattgaact caggacctct ggaagagcag gcagtgttct taaccgcgtg acatctttcc 360
agaccaagaa atcatttttt aaaaacacac acgtaacgta aaatgggaaa tgggactctc 420
tctgaggctg tggagctaga actcaggact ctttgacac tatgacctct gctcactgt 480
ggccaacgtt tcagaaaaag catttgggtg agacttgggg ttctactggg gagcgtcctg 540
gggacagtgg gtcccaatgg ctctagggtc ctctgtgta gctggggaga taacacaaaa 600
gggattcttt tgaggcttcc aacaggagat aggacctggt gagcctttgt ctctctgcat 660

2

cgt aag ctg ctg gtg ctg ggc acc tgg ttg ctg gca ctg ctg ctg acc	1363
Arg Lys Leu Leu Val Leu Gly Thr Trp Leu Leu Ala Leu Leu Leu Thr	
170 175 180	
cta ccc atg atg ctt gcc atc cag ctg gtc cgc agg ggc tct aag agc	1411
Leu Pro Met Met Leu Ala Ile Gln Leu Val Arg Arg Gly Ser Lys Ser	
185 190 195	
ctc tgc ctg cca gcc tgg ggc cct cgt gcc cac cgt act tac cta acg	1459
Leu Cys Leu Pro Ala Trp Gly Pro Arg Ala His Arg Thr Tyr Leu Thr	
200 205 210	
ttg ctc ttt ggg acc agc att gtg ggg cct ggc ttg gtc att ggg ctg	1507
Leu Leu Phe Gly Thr Ser Ile Val Gly Pro Gly Leu Val Ile Gly Leu	
215 220 225	
ctc tat gtc cgt ctg gcc agg gcc tac tgg cta tct cag caa gct tct	1555
Leu Tyr Val Arg Leu Ala Arg Ala Tyr Trp Leu Ser Gln Gln Ala Ser	
230 235 240 245	
ttc aag cag aca cgg cgg ctg ccc aac ccc agg gtg ctc tac ctc atc	1603
Phe Lys Gln Thr Arg Arg Leu Pro Asn Pro Arg Val Leu Tyr Leu Ile	
250 255 260	
ctt ggt atc gtc ctt ctc ttc tgg gcc tgc ttt cta ccc ttc tgg ctg	1651
Leu Gly Ile Val Leu Leu Phe Trp Ala Cys Phe Leu Pro Phe Trp Leu	
265 270 275	
tgg cag ctg ctg gcc cag tac cac gag gcc atg cca ctg act ccc gag	1699
Trp Gln Leu Leu Ala Gln Tyr His Glu Ala Met Pro Leu Thr Pro Glu	
280 285 290	
act gca cgc att gtc aac tac ctg acc acc tgc ctc act tat ggc aac	1747
Thr Ala Arg Ile Val Asn Tyr Leu Thr Thr Cys Leu Thr Tyr Gly Asn	
295 300 305	
agt tgc atc aat ccc ttc ctc tac act ctg ctc acc aag aac tat cga	1795
Ser Cys Ile Asn Pro Phe Leu Tyr Thr Leu Leu Thr Lys Asn Tyr Arg	
310 315 320 325	
gag tac cta cgt ggc cgc cag cgg tca ctg ggt agt agt tgc cac agc	1843
Glu Tyr Leu Arg Gly Arg Gln Arg Ser Leu Gly Ser Ser Cys His Ser	
330 335 340	
cca ggg agt cct ggc agc ttc ctg ccc agc cga gtc cac ctc cag cag	1891
Pro Gly Ser Pro Gly Ser Phe Leu Pro Ser Arg Val His Leu Gln Gln	
345 350 355	

gac tcg ggc cgc tcg ctg tcc tcc agc agc caa cag gcc aca gag acc 1939
 Asp Ser Gly Arg Ser Leu Ser Ser Ser Ser Gln Gln Ala Thr Glu Thr
 360 365 370

ctc atg ctg tct cca gtc ccc cgt aac ggg gcc ctt ctc tga 1981
 Leu Met Leu Ser Pro Val Pro Arg Asn Gly Ala Leu Leu
 375 380 385

gagtgcactg tgcaatcctg gcataggaaa ggacccaaag gcgtgcggct ccggagcgca 2041
 ttcccagaat cccctgctca aacctaactg gctcgtctgt cttcttcag ttctcttctg 2101
 gaaagccccc tgttgtttct tcgttgctca cttcagctct ttccggcaact ttgttacgat 2161
 gccaacacat acacaaccaa tctctyaggg acttcggaag ctgtggtgtg ccgagtcac 2221
 taaggcctaa ttgccagag cagctagtagc aggctgacct gtgatgtaag cccttcacagc 2281
 cctgaggacc ttcccgact ggtggttaagg gaccggcagg gtatctctac cctctcccca 2341
 tcccatggta gcttcttgct tccgtatgca catggaacaa taaacatgt actctgcagt 2401
 gtgtcctgcc tcagtgttg ggatggaagt tgtggtggg tggaatcct gaactctgcg 2461
 atccccagca tgaggaaact aagtgtttg tgaggtttcc aaacctttgc ctgtggatgt 2521
 ccctacatgg ttgtccttg gacagctttt ccttggtga ctgtttctct taccacgggt 2581
 aaaagacagc tgcccacct cttcaggggtg accatcagag ggtgggacgg cacctgggtc 2641
 tgtagtact cttgctgaat gtggcagggc atggagtga ggtctctgca gcgacaaagt 2701
 gttccagaag tagattctgg tgtctgacca cagtactttg cttttttatt ttgctattaa 2761
 aaatgagaga taccatacc cacatgcgca cacacgcaga gagagacaca cgcagagaga 2821
 gacacacgca gagagaaatg aaggggttta gcctgtaatt ccaggagttg gggaggcttg 2881
 aggtcagcct gtgctactta ttaaagactt taaatttaaa aaaaaaaca gggttgggga 2941
 tttagctcag tggtagagcg cttgcctaga aagcgcaagg ccctgggttc ggtcctcagc 3001
 tccgaaaaaa agaaaaagaa aaacaaacaa acaaactttt ctttgagaca gggctctcact 3061
 atgcagcctt ggctggcctt agacttactg tgtagacaag gctggcctga tctcacagag 3121
 a 3122

<210> 2

<211> 386

<212> PRT

<213> Rat

<400> 2

Met Ala Leu Ser Leu Glu Ser Thr Thr Ser Phe His Met Leu Thr Val
 1 5 10 15

Ser Gly Ser Thr Val Thr Glu Leu Pro Gly Asp Ser Asn Val Ser Leu
 20 25 30

Asn Ser Ser Trp Ser Gly Pro Thr Asp Pro Ser Ser Leu Lys Asp Leu
 35 40 45

Val Ala Thr Gly Val Ile Gly Ala Val Leu Ser Ala Met Gly Val Val
 50 55 60

Gly Met Val Gly Asn Val Tyr Thr Leu Val Val Met Cys Arg Phe Leu
 65 70 75 80

Arg Ala Ser Ala Ser Met Tyr Val Tyr Val Val Asn Leu Ala Leu Ala
 85 90 95

Asp Leu Leu Tyr Leu Leu Ser Ile Pro Phe Ile Ile Ala Thr Tyr Val
 100 105 110

Thr Lys Asp Trp His Phe Gly Asp Val Gly Cys Arg Val Leu Phe Ser
 115 120 125

Leu Asp Phe Leu Thr Met His Ala Ser Ile Phe Thr Leu Thr Ile Met
 130 135 140

Ser Ser Glu Arg Tyr Ala Ala Val Leu Arg Pro Leu Asp Thr Val Gln
 145 150 155 160

Arg Ser Lys Gly Tyr Arg Lys Leu Leu Val Leu Gly Thr Trp Leu Leu
 165 170 175

Ala Leu Leu Leu Thr Leu Pro Met Met Leu Ala Ile Gln Leu Val Arg
 180 185 190

Arg Gly Ser Lys Ser Leu Cys Leu Pro Ala Trp Gly Pro Arg Ala His
 195 200 205

Arg Thr Tyr Leu Thr Leu Leu Phe Gly Thr Ser Ile Val Gly Pro Gly
 210 215 220

Leu Val Ile Gly Leu Leu Tyr Val Arg Leu Ala Arg Ala Tyr Trp Leu
 225 230 235 240

Ser Gln Gln Ala Ser Phe Lys Gln Thr Arg Arg Leu Pro Asn Pro Arg
 245 250 255

Val Leu Tyr Leu Ile Leu Gly Ile Val Leu Leu Phe Trp Ala Cys Phe
 260 265 270

Leu Pro Phe Trp Leu Trp Gln Leu Leu Ala Gln Tyr His Glu Ala Met
 275 280 285

Pro Leu Thr Pro Glu Thr Ala Arg Ile Val Asn Tyr Leu Thr Thr Cys
 290 295 300

Leu Thr Tyr Gly Asn Ser Cys Ile Asn Pro Phe Leu Tyr Thr Leu Leu
 305 310 315 320

Thr Lys Asn Tyr Arg Glu Tyr Leu Arg Gly Arg Gln Arg Ser Leu Gly
 325 330 335

Ser Ser Cys His Ser Pro Gly Ser Pro Gly Ser Phe Leu Pro Ser Arg
 340 345 350

Val His Leu Gln Gln Asp Ser Gly Arg Ser Leu Ser Ser Ser Ser Gln
 355 360 365

Gln Ala Thr Glu Thr Leu Met Leu Ser Pro Val Pro Arg Asn Gly Ala
 370 375 380

Leu Leu
 385

<210> 3

<211> 6

<212> PRT

<213> Homo sapiens

<220>

<221> DISULFID

<222> (1)..(6)

<223> Cyclic core

<400> 3

Cys Phe Trp Lys Tyr Cys
 1 5

<210> 4
<211> 12
<212> PRT
<213> Gobius sp.

<220>
<221> DISULFID
<222> (6)..(11)
<223> cyclic core

<400> 4
Ala Gly Thr Ala Asp Cys Phe Trp Lys Tyr Cys Val
1 5 10

<210> 5
<211> 11
<212> PRT
<213> Homo sapiens

<220>
<221> DISULFID
<222> (5)..(10)
<223> Cyclic core

<400> 5
Glu Thr Pro Asp Cys Phe Trp Lys Tyr Cys Val
1 5 10

<210> 6
<211> 6
<212> PRT
<213> Artificial Sequence

<220>
<221> DISULFID
<222> (1)..(6)
<223> Cyclic core

<220>
<221> MOD_RES
<222> (1)
<223> ACETYLTATION

<220>

<221> SITE
<222> (3)
<223> D-Trp

<220>
<221> MOD_RES
<222> (6)
<223> AMIDATION

<220>
<223> Description of Artificial Sequence: synthetic
construct

<400> 6
Cys Phe Trp Lys Tyr Cys
1 5

<210> 7
<211> 14
<212> PRT
<213> Homo sapiens

<220>
<221> DISULFID
<222> (3)..(14)

<400> 7
Ala Gly Cys Lys Asn Phe Phe Trp Lys Thr Phe Thr Ser Cys
1 5 10

<210> 8
<211> 14
<212> PRT
<213> Homo sapiens

<220>
<221> DISULFID
<222> (2)..(13)

<400> 8
Pro Cys Lys Asn Phe Phe Trp Lys Thr Phe Ser Ser Cys Lys
1 5 10

<210> 9
<211> 8

<212> PRT

<213> Artificial Sequence

<220>

<221> SITE

<222> (1)

<223> D-Phe

<220>

<221> SITE

<222> (4)

<223> D-Trp

<220>

<221> DISULFID

<222> (2)..(7)

<220>

<221> MOD_RES

<222> (8)

<223> AMIDATION

<220>

<223> Description of Artificial Sequence: synthetic
construct

<400> 9

Phe Cys Tyr Trp Lys Val Cys Trp

1 5

<210> 10

<211> 8

<212> PRT

<213> Artificial Sequence

<220>

<221> SITE

<222> (1)

<223> D-Phe

<220>

<221> SITE

<222> (4)

<223> D-Trp

<220>

<221> DISULFID

<222> (2) .. (7)

<220>

<223> Description of Artificial Sequence: synthetic construct

<400> 10

Phe Cys Phe Trp Lys Thr Cys Thr
1 5

<210> 11

<211> 24

<212> DNA

<213> Homo sapiens

<400> 11

ctgagcctgg agtctacaac aagc 24

<210> 12

<211> 24

<212> DNA

<213> Homo sapiens

<400> 12

taggtggcta tgatgaaggg aatg 24

INTERNATIONAL SEARCH REPORT

Intern Application No

PCT/US 00/21171

A. CLASSIFICATION OF SUBJECT MATTER

IPC 7 G01N33/74 C07K14/72 C07K14/575 C07K14/705

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

IPC 7 G01N

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

EPO-Internal, WPI Data, PAJ, BIOSIS, MEDLINE

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	WO 99 40192 A (SMITHKLINE BEECHAM LAB ;SMITHKLINE BEECHAM PLC (GB); SMITHKLINE BE) 12 August 1999 (1999-08-12) abstract; claims 1,23-33 examples 3,4,7-9	1-17
X	WO 99 35266 A (SMITHKLINE BEECHAM LAB ;SMITHKLINE BEECHAM PLC (GB); SMITHKLINE BE) 15 July 1999 (1999-07-15) page 17, line 23 -page 18, line 11; claim 6	1,3-17
X	WO 95 34651 A (CHALMERS DEREK ;NEUROCRINE BIOSCIENCES INC (US)) 21 December 1995 (1995-12-21) the whole document	1,3-17
-/--		



Further documents are listed in the continuation of box C.



Patent family members are listed in annex.

* Special categories of cited documents :

- *A* document defining the general state of the art which is not considered to be of particular relevance
- *E* earlier document but published on or after the international filing date
- *L* document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)
- *O* document referring to an oral disclosure, use, exhibition or other means
- *P* document published prior to the international filing date but later than the priority date claimed

- *T* later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
- *X* document of particular relevance: the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
- *Y* document of particular relevance: the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.
- *G* document member of the same patent family

Date of the actual completion of the international search

18 January 2001

Date of mailing of the international search report

26/01/2001

Name and mailing address of the ISA

European Patent Office, P.B. 5818 Patentlaan 2
NL - 2280 HV Rijswijk
Tel. (+31-70) 340-2040. Tx. 31 651 epo nl,
Fax (+31-70) 340-3016

Authorized officer

Gundlach, B

INTERNATIONAL SEARCH REPORT

International Application No

PCT/US 00/21171

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	<p>GRIGORIADIS DIMITRI E ET AL: "125I-Tyr-0-sauvagine: A novel high affinity radioligand for the pharmacological and biochemical study of human corticotropin-releasing factor-2alpha receptors." MOLECULAR PHARMACOLOGY, vol. 50, no. 3, September 1996 (1996-09), pages 679-686, XP000978435 ISSN: 0026-895X the whole document</p>	1,4,5
X	<p>LUTHIN DAVID R ET AL: "Synthesis and biological activity of oxo-7H-benzo(e)perimidine-4-carb oxylic acid derivatives as potent, nonpeptide Corticotropin Releasing Factor (CRF) receptor antagonists." BIOORGANIC & MEDICINAL CHEMISTRY LETTERS, vol. 9, no. 5, 8 March 1999 (1999-03-08), pages 765-770, XP000978428 ISSN: 0960-894X abstract</p>	1,4,5
X	<p>GIBSON A: "COMPLEX EFFECTS OF GILlichTHYS UROTENSIN II ON RAT AORTIC STRIPS" BRITISH JOURNAL OF PHARMACOLOGY, vol. 91, no. 1, 1987, pages 205-212, XP000978590 ISSN: 0007-1188 abstract; table 1</p>	1,4
X	<p>ARDATI ALI ET AL: "Human CRF2 alpha and beta splice variants: Pharmacological characterization using radioligand binding and a luciferase gene expression assay." NEUROPHARMACOLOGY, vol. 38, no. 3, March 1999 (1999-03), pages 441-448, XP000978443 ISSN: 0028-3908 the whole document</p>	1,4,5
A	<p>EP 0 859 052 A (SMITHKLINE BEECHAM CORP) 19 August 1998 (1998-08-19) the whole document</p>	1-17
P,A	<p>NOTHACKER H P ET AL: "Identification of the natural ligand of an orphan G-protein-coupled receptor involved in the regulation of vasoconstriction." NAT CELL BIOL, (1999 OCT) 1 (6) 383-5. , XP000980244 the whole document</p>	1-17

INTERNATIONAL SEARCH REPORT

Information on patent family members

International Application No

PCT/US 00/21171

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
WO 9940192 A	12-08-1999	US 6133420 A US 6159700 A EP 1056844 A	17-10-2000 12-12-2000 06-12-2000
WO 9935266 A	15-07-1999	US 6075137 A EP 1045908 A	13-06-2000 25-10-2000
WO 9534651 A	21-12-1995	US 5786203 A AT 171212 T AU 703140 B AU 2945595 A CA 2193072 A DE 69504808 D DE 69504808 T DK 724637 T EP 0724637 A EP 0860501 A ES 2123269 T JP 10501420 T	28-07-1998 15-10-1998 18-03-1999 05-01-1996 21-12-1995 22-10-1998 12-05-1999 14-06-1999 07-08-1996 26-08-1998 01-01-1999 10-02-1998
EP 0859052 A	19-08-1998	US 5851798 A JP 10295376 A US 6133420 A US 6005074 A US 6159700 A	22-12-1998 10-11-1998 17-10-2000 21-12-1999 12-12-2000

PCTWORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification : C12N	A2	(11) International Publication Number: WO 00/02996 (43) International Publication Date: 20 January 2000 (20.01.00)
<p>(21) International Application Number: PCT/US99/15425</p> <p>(22) International Filing Date: 9 July 1999 (09.07.99)</p> <p>(30) Priority Data: 60/092,357 10 July 1998 (10.07.98) US</p> <p>(71) Applicant: CORNELL RESEARCH FOUNDATION, INC. [US/US]; Suite 105, 20 Thornwood Drive, Ithaca, NY 14850 (US).</p> <p>(72) Inventors: BAUER, David, W.; 10815 115th Court N.E., Kirkland, WA 98033 (US). BEER, Steven, V.; 211 Hudson Street, Ithaca, NY 14850 (US). BOGDANOVE, Adam, J.; 37 Uptown Road, Ithaca, NY 14850 (US). COLLMER, Alan; 139 Lexington Drive, Ithaca, NY 14850 (US). HAM, Jong, Hyun; 17-D Gaslight Village Apartments, Ithaca, NY 14850 (US).</p> <p>(74) Agents: GOLDMAN, Michael, L. et al.; Nixon Peabody LLP, Clinton Square, P.O. Box 1051, Rochester, NY 14603 (US).</p>		<p>(81) Designated States: AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE, DK, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, UA, UG, UZ, VN, YU, ZW, ARIPO patent (GH, GM, KE, LS, MW, SD, SL, SZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).</p> <p>Published <i>Without international search report and to be republished upon receipt of that report.</i></p>
<p>(54) Title: RECOMBINANT CONSTRUCTS AND SYSTEMS FOR SECRETION OF PROTEINS VIA TYPE III SECRETION SYSTEMS</p>		
<p>(57) Abstract</p> <p>One aspect of the present invention relates to a DNA construct that contains a first DNA molecule encoding a functional type III secretion system, a promoter, and a second DNA molecule encoding a protein or polypeptide capable of being secreted by the type III secretion system. The second DNA molecule is operably coupled to the promoter so that upon introduction of the DNA construct into a host cell, the encoded protein or polypeptide and the type III secretion system are expressed and the encoded protein or polypeptide is secreted. Another aspect of the present invention relates to a system that includes (i) a first DNA construct having a first DNA molecule encoding a functional type III secretion system and (ii) a second DNA construct having a promoter operably coupled to a second DNA molecule encoding a protein or polypeptide capable of being secreted by the type III secretion system. Upon introduction of the first and second DNA constructs into a host cell, the encoded protein or polypeptide and the type III secretion system are expressed and the encoded protein or polypeptide is secreted. Methods of isolating a protein or polypeptide and identifying a gene encoding a potential effector protein or polypeptide are also disclosed.</p>		

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	ES	Spain	LS	Lesotho	SI	Slovenia
AM	Armenia	FI	Finland	LT	Lithuania	SK	Slovakia
AT	Austria	FR	France	LU	Luxembourg	SN	Senegal
AU	Australia	GA	Gabon	LV	Latvia	SZ	Swaziland
AZ	Azerbaijan	GB	United Kingdom	MC	Monaco	TD	Chad
BA	Bosnia and Herzegovina	GE	Georgia	MD	Republic of Moldova	TG	Togo
BB	Barbados	GH	Ghana	MG	Madagascar	TJ	Tajikistan
BE	Belgium	GN	Guinea	MK	The former Yugoslav Republic of Macedonia	TM	Turkmenistan
BF	Burkina Faso	GR	Greece	ML	Mali	TR	Turkey
BG	Bulgaria	HU	Hungary	MN	Mongolia	TT	Trinidad and Tobago
BJ	Benin	IE	Ireland	MR	Mauritania	UA	Ukraine
BR	Brazil	IL	Israel	MW	Malawi	UG	Uganda
BY	Belarus	IS	Iceland	MX	Mexico	US	United States of America
CA	Canada	IT	Italy	NE	Niger	UZ	Uzbekistan
CF	Central African Republic	JP	Japan	NL	Netherlands	VN	Viet Nam
CG	Congo	KE	Kenya	NO	Norway	YU	Yugoslavia
CH	Switzerland	KG	Kyrgyzstan	NZ	New Zealand	ZW	Zimbabwe
CI	Côte d'Ivoire	KP	Democratic People's Republic of Korea	PL	Poland		
CM	Cameroon	KR	Republic of Korea	PT	Portugal		
CN	China	KZ	Kazakhstan	RO	Romania		
CU	Cuba	LC	Saint Lucia	RU	Russian Federation		
CZ	Czech Republic	LI	Liechtenstein	SD	Sudan		
DE	Germany	LK	Sri Lanka	SE	Sweden		
DK	Denmark	LR	Liberia	SG	Singapore		
EE	Estonia						

RECOMBINANT CONSTRUCTS AND SYSTEMS FOR SECRETION OF PROTEINS VIA TYPE III SECRETION SYSTEMS

This application claims the benefit of U.S. Provisional Patent Application
5 Serial No. 60/092,357, filed July 10, 1998.

This invention was made in part with support by the U.S. Government
under Grant No. MCB-9631530 from National Science Foundation and Grant No. 97-
35303-4488 from the U.S. Department of Agriculture, NRI Competitive Grants Program.
The U.S. Government may have certain rights in this invention.

10

BACKGROUND OF THE INVENTION

The most common bacterial pathogens of plants colonize the apoplast,
and from that location outside of the walls of living cells they incite a variety of diseases
15 in most cultivated plants (Alfano et al., "Bacterial Pathogens in Plants: Life Up Against
the Wall," Plant Cell 8:1683-1698 (1996)). The majority of these are Gram-negative
bacteria in the genera *Erwinia*, *Pseudomonas*, *Xanthomonas*, and *Ralstonia*. Most are
host specific and will elicit the hypersensitive response ("HR") in nonhosts. The HR is a
rapid, programmed death of plant cells in contact with the pathogen. Some of the
20 defense responses associated with the HR are localized at the periphery of plant cells at
the site of bacterial contact, but what actually stops bacterial growth is not known
(Brown et al., "hrp genes in *Xanthomonas campestris* pv. *vesicatoria* Determine Ability
to Suppress Papilla Deposition in Pepper Mesophyll Cells," MPMI 8:825-836 (1995);
Young et al., "Changes in the Plasma Membrane Distribution of Rice Phospholipase D
25 During Resistant Interactions With *Xanthomonas oryzae* pv. *oryzae*," Plant Cell 8:1079-
1090 (1996); Bestwick et al., "Localization of Hydrogen Peroxide Accumulation During
the Hypersensitive Reaction of Lettuce Cells to *Pseudomonas syringae* pv.
phaseolicola," Plant Cell 9:209-221 (1997)). Pathogenesis in host plants, in contrast,
involves prolonged bacterial multiplication, spread to surrounding tissues, and the
30 eventual production of macroscopic symptoms characteristic of the disease. Although
these bacteria are diverse in their taxonomy and pathology, they all possess *hrp* genes
which direct their ability to elicit the HR in nonhosts or to be pathogenic (and parasitic)
in hosts (Lindgren, "The Role of *hrp* Genes During Plant-Bacterial Interactions," Annu.
Rev. Phytopathol. 35:129-152 (1997)). The *hrp* genes encode a type III protein secretion

system that appears to be capable of delivering Avr (avirulence) proteins across the walls and plasma membranes of living plant cells (Alfano et al., "The Type III (Hrp) Secretion Pathway of Plant Pathogenic Bacteria: Trafficking Harpins, Avr Proteins, and Death," J. Bacteriol. 179:5655-5662 (1997), which is hereby incorporated by reference). The Avr proteins are so named because they can betray the parasite to the *R* gene-encoded surveillance system of plants, thereby triggering the HR (Vivian et al., "Avirulence Genes in Plant-Pathogenic Bacteria: Signals or Weapons?," Microbiology 143:693-704 (1997); Leach et al., "Bacterial Avirulence Genes," Annul. Rev. Phytopathol. 34:153-179 (1996)). But Avr-like proteins also appear to be key to parasitism in compatible host plants, where the parasite proteins are undetected and the HR is not triggered. Thus, bacterial avirulence and pathogenicity are interrelated phenomena and explorations of HR elicitation are furthering our understanding of parasitic mechanisms.

Despite the emerging importance of Avr proteins, there is no direct evidence that they travel the Hrp pathway, there is no knowledge of their function in virulence, it appears likely that only a subset of those that are produced by typical host-specific pathogens have been identified, and there is no evidence that they are produced at all by host-promiscuous pathogens. The evidence that Avr proteins are transferred by the Hrp pathway into plants is most complete, although still indirect, with *Pseudomonas syringae* AvrB and AvrPto proteins. Nonpathogenic *Escherichia coli* and *Pseudomonas fluorescens* cells that harbor the functional cluster of *Pseudomonas syringae* *hrp* genes carried on cosmid pHIR11 can elicit an HR that is dependent on both the type III secretion system and either AvrB or AvrPto (Gopalan et al., "Expression of the *Pseudomonas Syringae* Avirulence Protein AvrB in Plant Cells Alleviates its Dependence on the Hypersensitive Response and pathogenicity (Hrp) Secretion System in Eliciting Genotype-specific Hypersensitive Cell Death," Plant Cell 8:1095-1105 (1996); Pirhonen et al., "Phenotypic Expression of *Pseudomonas Syringae* *avr* Genes in *E. Coli* is Linked to the Activities of the *hrp*-encoded Secretion System," MPMI 9:252-260 (1996)). Both Avr proteins trigger an *R* gene-dependent HR when transiently expressed inside plant cells (Gopalan et al., "Expression of the *Pseudomonas Syringae* Avirulence Protein AvrB in Plant Cells Alleviates its Dependence on the Hypersensitive Response and pathogenicity (Hrp) Secretion System in Eliciting Genotype-specific Hypersensitive Cell Death," Plant Cell 8:1095-1105 (1996)) and the interaction of AvrPto and Pto in the yeast two-hybrid system correlates with biological activity (Tang

et al., Science 274:2060 (1996); Scofield et al., Science 274:2063-2065 (1996)).
However, neither *Pseudomonas syringae*, *Escherichia coli* (pHIR11), nor
Pseudomonas fluorescens (pHIR11) secrete AvrB or AvrPto in culture, presumably
because these proteins travel the type III pathway directly into host cells and only upon
5 host cell contact, as with the Yop virulence proteins of *Yersinia* spp. (Gopalan et al.,
"Expression of the *Pseudomonas syringae* Avirulence Protein AvrB in Plant Cells
Alleviates its Dependence on the Hypersensitive Response and Pathogenicity (Hrp)
Secretion System in Eliciting Genotype-specific Hypersensitive Cell Death," Plant Cell
8:1095-1105 (1996); Cornelis et al., "The *Yersinia* Yop Regulon: A Bacterial System for
10 Subverting Eukaryotic Cells," Mol. Microbiol. 23:861-867 (1997)). Other known Avr
proteins have been observed only in the bacterial cytoplasm (Leach et al., "Bacterial
Avirulence Genes," Annu. Rev. Phytopathol. 34:153-179 (1996); Knoop et al.,
"Expression of the Avirulence Gene avrBs3 from *Xanthomonas campestris* pv.
vesicatoria is not Under the Control of hrp Genes and is Independent of Plant Factors," J.
15 Bacteriol. 173:7142-7150 (1991); Puri et al., "Expression of *avrPphB*, an Avirulence
Gene from *Pseudomonas Syringae* pv. *Phaseolicola*, and the Delivery of Signals
Causing the Hypersensitive Reaction in Bean," MPMI 10:247-256 (1997)).

Many proteins and polypeptides, including hormones and enzymes, are in
high demand for pharmacological and industrial use. Once the gene encoding a desired
20 protein or polypeptide has been isolated, the protein can be produced readily through
fermentation in rapidly growing bacteria. *Escherichia coli* is used most commonly for
large-scale protein production. Current technology enables the production of relatively
large intracellular concentrations of the desired proteins or polypeptides. Extraction of
the desired protein or polypeptide from the bacterial cells requires lysing of the cell
25 membrane. After lysing the cell membrane, the desired protein or polypeptide is
contaminated with other proteins and, therefore, subject to degradation. The resulting
contamination requires significant purification to obtain the isolated protein or
polypeptide and degradation of the desired protein or polypeptide limits the obtainable
yield.

30 In addition to fermentation technologies for production of proteins or
polypeptides, gene therapy involving transgenic plants is emerging as an important tool
for enhancing agricultural productivity and reducing disease losses. For example,
transgenic plants expressing bacterial and viral proteins are now used for herbicide

Thus, it would be beneficial to obtain a recombinant construct and expression system which overcomes these and other deficiencies in the art, particularly the ability to produce a recombinant host organism capable of expressing and secreting Avr and/or other desired proteins or polypeptides into their environment (i.e., culture medium).

One aspect of the present invention relates to a DNA construct that contains a first DNA molecule encoding a functional type III secretion system, a promoter, and a second DNA molecule encoding a protein or polypeptide capable of being secreted by the type III secretion system. The second DNA molecule is operably coupled to the promoter so that upon introduction of the DNA construct into a host cell, the encoded protein or polypeptide and the type III secretion system are expressed and the encoded protein or polypeptide is secreted. Also disclosed are host cells and expression systems that contain the DNA construct, as well as a method of secreting a protein or polypeptide into the environment of a host cell which employs the DNA construct.

Another aspect of the present invention relates to a system that includes a

(i) first DNA construct having a first DNA molecule encoding a functional type III

secretion system and (ii) a second DNA construct having a promoter operably coupled to a second DNA molecule encoding a protein or polypeptide capable of being secreted by the type III secretion system. Upon introduction of the first and second DNA constructs into a host cell, the encoded protein or polypeptide and the type III secretion system are
5 expressed and the encoded protein or polypeptide is secreted. Also disclosed are host cells and expression systems that contain the system of DNA constructs, as well as a method of secreting a protein or polypeptide into the environment of a host cell which employs the system of DNA constructs.

A further aspect of the present invention relates to a method of isolating a
10 protein or polypeptide. This method is performed by providing a recombinant host cell that contains (i) a first DNA molecule encoding a functional type III secretion system and (ii) a second, heterologous DNA molecule having a promoter operably coupled to a nucleic acid sequence encoding a protein or polypeptide capable of being secreted by the type III secretion system. The recombinant host cell is introduced into a culture medium,
15 wherein the encoded protein or polypeptide and the type III secretion system are expressed and the encoded protein or polypeptide is secreted into the culture medium. Subsequently, the encoded protein or polypeptide is isolated from the culture medium.

Still another aspect of the present invention relates to a method of
identifying a gene encoding a potential effector protein or polypeptide. This method of
20 the invention is performed by providing a host cell that contains a DNA molecule encoding a functional type III secretion system. Next, a candidate gene encoding a protein or polypeptide is inserted into the host cell under conditions effective to express the encoded protein or polypeptide. Finally, it is determined whether the encoded protein or polypeptide is secreted by the recombinant host cell, wherein secretion of the
25 encoded protein or polypeptide indicates that the gene encodes a potential effector protein or polypeptide.

Since the DNA constructs of the present invention enable expression and secretion of proteins by recombinant host cells, it is possible to employ these recombinant host cells in a fermentation system which enables efficient production of a
30 desired protein or polypeptide that can be purified at high yield and at minimal expense compared to existing fermentation/purification procedures. Moreover, the constructs of the present invention can be employed to bioprospect for potential effector proteins or polypeptides, which by virtue of their expression and secretion by a recombinant host

cell expressing a type III secretion system, become likely candidates as effector proteins. This method of screening for potential effector protein is novel and much more systematic and efficient than prior methods.

5 BRIEF DESCRIPTION OF THE DRAWINGS

Figure 1 is a diagram of the physical maps for cosmids pCPP2156 and pCPP2157, which contain the *Erwinia chrysanthemi* *hrp* region, and comparison of the *hrp* regions of *Erwinia chrysanthemi* and *Erwinia amylovora* (Bauer et al., "*Erwinia chrysanthemi* harpin_{Ech}: An Elicitor of the Hypersensitive Response that Contributes to Soft-rot Pathogenesis," MPMI 8:484-491 (1995); Kim et al., "The *hrpC* and *hrpN* Operons of *Erwinia chrysanthemi* EC16 are Flanked by *plcA* and Homologs of Hemolysin/Adhesin Genes and Accompanying Activator/Transporter Genes," MPMI 11(6):563-567 (1998); Bogdanove et al., "*Erwinia amylovora* Secretes Harpin via a Type III of Pathway and Contains a Homolog of YopN of *Yersinia* spp.," J. Bacteriol. 178:1720-1730 (1996); Wei et al., "Harpin, Elicitor of the Hypersensitive Response Produced by the Plant Pathogen *Erwinia amylovora*," Science, 257:85-88 (1992); Wei et al., "HrpI of *Erwinia amylovora* Functions in Secretion of Harpin and is a Member of a New Protein Family," J. Bacteriol., 175:7958-7967 (1993); Kim et al., "The HrpA and HrpC Operons of *Erwinia amylovora* Encode Components of a Type III Pathway that Secretes Harpin," J. Bacteriol. 179:1690-1697 (1997), which are hereby incorporated by reference). Arrow-shaped boxes denote putative transcriptional units. Shadowed areas denote *hrp* regions. Dashed boxes denote transcriptional units predicted on the basis of the homology and spacing of partially sequenced regions (shaded areas) in comparison with the corresponding *Erwinia amylovora* *hrp* genes. The filled triangle indicates the location of mini-Tn5Cm in pCPP2368.

Figure 2A contains images of *Nicotiana clevelandii* leaves infiltrated with *Escherichia coli* DH5 α carrying *Erwinia chrysanthemi* *hrp* clusters that are either intact (pCPP2156, pCPP2416) or defective (pCPP2157, pCPP2368) at a concentration of 5 X 10⁸ cfu/ml. Leaves were photographed 48 hours after infiltration. Tissue collapse occurred within 24 hours. The areas below each number on the leaf was infiltrated with *Escherichia coli* DH5 α carrying following constructs: 1, pCPP2156; 2, pCPP2156 and

pAVRB-FLAG2; 3, pCPP2157; 4, PCPP2157 and pAVRB-FLAG2; 5, pCPP2416 and pAVRB-FLAG2; 6, pCPP2368 and pAVRB-FLAG2.

Figure 2B contains images of the leaves of two soybean cultivars, Norchief (*RPG1*) and Acme (*rpg1*), infiltrated with *Escherichia coli* DH5 α carrying *Erwinia chrysanthemi* *hrp* clusters that are either intact (pCPP2156, pCPP2416) or defective (pCPP2157, pCPP2368) at a concentration of 5×10^8 cfu/ml. Leaves were photographed 72 hours after infiltration. Tissue collapse occurred within 48 hours. The first two leaves (numbers 1 through 4) are Norchief and the third leaf (number 5 and 6) is Acme. The area below each number on the leaf was infiltrated with *Escherichia coli* DH5 α carrying following constructs: 1, pCPP2416; 2, pCPP2416 and pAVRB-FLAG2; 3, pCPP2368; 4, PCPP2368 and pAVRB-FLAG2; 5, pCPP2156; 6, pCPP2156 and pAVRB-FLAG2.

Figure 3 is an image of an immunodetection in western blots showing differential secretion of AvrB-FLAG by *Escherichia coli* DH5 α carrying either a wild type (pCPP2156) or mutant (pCPP2368) *Erwinia chrysanthemi* *hrp* cluster. The supernatant fraction (S) was concentrated 7.5 X more than the cell pellet fraction (P). Lanes: 1, *Escherichia coli*(pCPP2156), pAVRB-FLAG2; 2, *Escherichia coli*(pCPP2368, pAVRB-FLAG2); 3, *Escherichia coli*(pCPP2156, pAVRB-FLAG2); 4, *Escherichia coli*(pCPP2368), pAVRB-FLAG2).

Figure 4 is an image of an immunodetection in western blots showing differential secretion of AvrB by *Escherichia coli* DH5 α carrying either a wild type (pCPP2156) or mutant (pCPP2368) *Erwinia chrysanthemi* *hrp* cluster. pCPP2138 encodes mature β -lactamase, which was used as a cytoplasmic marker. The supernatant fraction (S) was concentrated 7.5 X more than the cell pellet fraction (P) for lanes 1 and 2 and 15 X more for lanes 6 and 7. Lanes: 1, *Escherichia coli*(pCPP2156, pAVRB1); 2, *Escherichia coli*(pCPP2368, pAVRB1); 3, *Escherichia coli*(pCPP2156, pAVRB1); 4, *Escherichia coli*(pCPP2368), pAVRB1; 5, purified AvrB; 6, *Escherichia coli*(pCPP2156, pAVRB1, pCPP2318); 7, *Escherichia coli*(pCPP2368, pAVRB1, pCPP2318); 8 *Escherichia coli*(pCPP2156, pAVRB1, pCPP2318); 9, *Escherichia coli*(pCPP2368, pAVRB1, pCRPP2318).

Figure 5 is an image of an immunodetection in western blots showing differential secretion of AvrPto-FLAG by *Escherichia coli* DH5 α carrying either intact (pCPP2156) or defective (pCPP2157, pCPP2368) *Erwinia chrysanthemi* *hrp* clusters and

Escherichia coli MC4100 carrying pHIR11. As indicated above, pCPP2138 encodes mature β -lactamase, which was used as a cytoplasmic marker. The supernatant fraction (S) was concentrated 7.5 X more than the cell pellet fraction (P). Lanes: 1, *Escherichia coli*(pCPP2156), pCPP2318, pAVRPTO-FLAG); 2, *Escherichia coli*(pCPP2157, pCPP2318, pAVRPTO-FLAG); 3, *Escherichia coli*(pCPP2368, pCPP2318, pAVRPTO-FLAG); 4, *Escherichia coli*(pCPP2156, pCPP2318, pAVRPTO-FLAG); 5, *Escherichia coli*(pCPP2157, pCPP2318, pAVRPTO-FLAG); 6, *Escherichia coli*(pCPP2368, pCPP2318, pAVRPTO-FLAG); 7, *Escherichia coli*(pHIR11, pAVRPTO-FLAG); 8, *Escherichia coli*(pHIR11, pAVRPTO-FLAG).

Figure 6 is an image of the immunodetection of DspE and DspE Δ E1521 in western blots of culture supernatant and bacterial cell fractions using anti-DspE antiserum (see Bogdanove et al., *J. Bacteriol.* 180:2244-2247 (1998), which is hereby incorporated by reference). Lane 1, hrp mutant Ea273-K178; lane 2, wild-type strain Ea273; lane 3, partial dspE deletion mutant Ea273dspE Δ E1521. The migration of the molecular weight markers (BRL) is indicated at the left.

Figure 7 is a diagram showing the genetic organization of the hrp and dsp genes contained by cosmids pCPP430 and pCPP431. The letters designating the known or proposed functions correspond to the following: S, secretion; R, regulation; H, harpin; A, avirulence; D, disease; U, unknown.

Figure 8 is an image of an immunodetection in western blots of flagged *Pseudomonas syringae* Avr proteins in cell (C) and supernatant (S) fractions from cultures of *Escherichia coli* DH5 containing the designated Avr protein (AvrB-Flag or AvrPto-Flag) and the hrp/dsp cluster (pCPP430), the minimal hrp cluster (pCPP431), or the secretion-defective hrp cluster (pCPP430hrcV Δ).

DETAILED DESCRIPTION OF THE INVENTION

One aspect of the present invention relates to novel constructs that can be used to transform host cells so that they express and secrete (i.e., into the host cell environment) a protein or polypeptide of interest.

According to one embodiment, a DNA construct is provided which includes a first DNA molecule encoding a functional type III secretion system, a promoter, and a second DNA molecule encoding a protein or polypeptide capable of

being secreted by the type III secretion system. The second DNA molecule is operably coupled to the promoter so that upon introduction of the DNA construct into a host cell, the encoded protein or polypeptide and the type III secretion system are expressed and the encoded protein or polypeptide is secreted by the host cell.

5 According to a second embodiment, a pair of DNA constructs are utilized as part of a system. The first DNA construct includes a DNA molecule encoding a functional type III secretion system. The second DNA construct includes a promoter operably coupled to a DNA molecule encoding a protein or polypeptide capable of being secreted by the type III secretion system. Upon introduction of the first and second DNA
10 constructs into a host cell, the protein or polypeptide and the type III secretion system are expressed and the protein or polypeptide is secreted.

As used in each of the above embodiments, any functional type III secretion system can be employed. By functional, it is intended that the type III secretion system contain all required genes under appropriate transcriptional and/or
15 translational control such that the secretion system can secrete proteins or polypeptides that are capable of being secreted. Preferred type III secretion systems are those obtained from the genus *Erwinia*, more preferably the harpin secretion systems obtained from *Erwinia amylovora* or *Erwinia chrysanthemi*, and *Pseudomonas*, more preferably the harpin secretion systems obtained from *Pseudomonas syringae*. For example, the
20 harpin secretion system of *Erwinia amylovora* is present on cosmid pCPP430 (Beer et al., "The *hrp* Gene Cluster of *Erwinia amylovora*," in Advances in Molecular Genetics of Plant-Microbe Interactions, Proceedings of the 5th International Symposium on the Molecular Genetics of Plant-Microbe Interactions, Interlaken, Switzerland, September, 1990, pp. 53-60 (1991) which is hereby incorporated by reference) and the harpin
25 secretion system of *Erwinia carotovora* is present in cosmid pCPP2156 (Ham et al., "A Cloned *Erwinia chrysanthemi* Hrp (type III Protein Secretion) System Functions in *Escherichia coli* to Deliver *Pseudomonas syringae* Avr Signals to Plant Cells and to Secrete Avr Proteins in Culture," Proc. Natl. Acad. Sci. USA, 95(17): 10206-11 (1998), which is hereby incorporated by reference). A diagram of cosmid pCPP430 is shown at
30 Figure 7 and a diagram of cosmid pCPP2156 is shown at Figure 1.

Type III protein secretion systems are present in bacterial pathogens of both animals and plants, and are typified by the type III system of *Yersinia* spp. (Finlay et al., "Common Themes in Microbial Pathogenicity Revisited," Microbiol. Mol. Biol.

Rev., 61:136-169 (1997); Cornelis et al., "The *Yersinia* Yop Regulon: A Bacterial System for Subverting Eukaryotic Cells," Mol. Microbiol., 23:861-867 (1997), which are hereby incorporated by reference). These animal pathogens are primarily extracellular parasites, and their Yops (*Yersinia* outer proteins) are secreted and translocated directly into host cells in a contact-dependent manner (Cornelis et al., "The *Yersinia* Yop Regulon: A Bacterial System for Subverting Eukaryotic Cells," Mol. Microbiol., 23:861-867 (1997), which is hereby incorporated by reference). A similar host-contact dependency may operate in most plant pathogenic bacteria. Nine of the *hrp* genes are universal components of type III secretion systems, and these have been renamed *hrc* (HR and conserved) and given the last-letter designation of their *Yersinia* homolog (with the exception of *hrcV*) (Bogdanove et al., "Unified Nomenclature for Broadly Conserved *hrp* Genes of Phytopathogenic Bacteria," Mol. Microbiol., 20:681-683 (1996), which is hereby incorporated by reference). The Hrc proteins enable protein movement across the bacterial inner and outer membranes independently of the general protein export (Sec) pathway (Charkowski et al., "Altered Localization of HrpZ in *Pseudomonas syringae* pv. *syringae* *hrp* Mutants Suggests That Different Components of the Type III Secretion Pathway Control Protein Translocation Across the Inner and Outer Membranes of Gram-negative Bacteria," J. Bacteriol., 179:3866-3874 (1997), which is hereby incorporated by reference). In contrast to the Hrc proteins, the Hrp proteins may be peripheral components of the Hrp secretion system and are more likely to perform type III secretion functions that are extracellular and specific to protein transfer across the plant cell wall and plasma membrane.

The genes encoding type III secretion systems are usually clustered, and the emerging concept that genes with related functions in virulence are often grouped on plasmids or in horizontally-acquired pathogenicity islands has important implications throughout pathogenic microbiology (Lawrence et al., "Selfish Operons: Horizontal Transfer May Drive the Evolution of Gene Clusters," Genetics, 143:1843-1860 (1996); Groisman et al., "Pathogenicity Islands: Bacterial Evolution in Quantum Leaps," Cell, 87:791-794 (1996); Hacker et al., "Pathogenicity Islands of Virulent Bacteria: Structure, Function and Impact on Microbial Evolution," Mol. Microbiol., 23:1089-1097 (1997), which are hereby incorporated by reference). There is some evidence for horizontal acquisition of *hrp* gene clusters in plant pathogenic bacteria, and the *hrp* cluster in *Ralstonia solanacearum* is carried on a megaplasmid (Alfano et al., "Bacterial Pathogens

in Plants: Life Up Against the Wall," Plant Cell, 8:1683-1698 (1996), which is hereby incorporated by reference). The finding of a plasmid-borne *hrp* gene cluster in *Erwinia herbicola* pv. *gypsophila* suggests that virulence may be acquired readily by plant-associated bacteria (Nizan et al., "The Presence of *hrp* Genes on the Pathogenicity-associated Plasmid of the Tumorigenic Bacterium *Erwinia herbicola* pv. *gypsophila*," MPMI, 10:677-682 (1997), which is hereby incorporated by reference). *Erwinia herbicola* is a common epiphyte that is usually benign, but strains classified as *Erwinia herbicola* pv. *gypsophila* cause galls on gypsophila and elicit the HR in tobacco. A 150-kb plasmid carries phytohormone biosynthetic genes and *hrp* genes, and the latter are required both for gall formation and HR elicitation (Nizan et al., "The presence of *hrp* genes on the pathogenicity-associated Plasmid of the Tumorigenic Bacterium *Erwinia herbicola* pv. *gypsophila*," MPMI, 10:677-682 (1997), which is hereby incorporated by reference).

The clustering of genes with related function is also consistent with the ability of some cloned *hrp* clusters to enable nonpathogens like *Escherichia coli* to elicit the HR. This has been reported for cosmids pHIR11 from *Pseudomonas syringae* pv. *syringae*, pCPP430 from *Erwinia amylovora*, pPPY430 from *Pseudomonas syringae* pv. *phaseolicola*, and pCPP2156 from *Erwinia chrysanthemi* (Alfano et al., "Bacterial Pathogens in Plants: Life Up Against the Wall," Plant Cell, 8:1683-1698 (1996); Puri et al., "Expression of *avrPphB*, an Avirulence Gene from *Pseudomonas syringae* pv. *phaseolicola*, and the Delivery of Signals Causing the Hypersensitive Reaction in Bean," MPMI, 10:247-256 (1997); Ham et al., "A Cloned *Erwinia chrysanthemi* Hrp (type III Protein Secretion) System Functions in *Escherichia coli* to Deliver *Pseudomonas syringae* Avr Signals to Plant Cells and to Secrete Avr Proteins in Culture," Proc. Natl. Acad. Sci. USA, 95(17): 10206-11 (1998), which are hereby incorporated by reference). Although these cosmids support heterologous HR elicitation, they do not enable *Escherichia coli* to become pathogenic. The basis for HR elicitation is best understood with pHIR11. The cosmid carries a 25-kb set of *hrp* genes that is intact and functional, as revealed by DNA sequencing and the ability to direct secretion of the HrpZ harpin (Alfano et al., "The Type III (Hrp) Secretion Pathway of Plant Pathogenic Bacteria: Trafficking Harpins, Avr Proteins, and Death," J. Bacteriol. 179:5655-5662 (1997), which is hereby incorporated by reference). The cosmid also carries, adjacent to the *hrp* cluster, the *hrmA* gene, which is *avr*-like in producing an avirulence phenotype when

expressed in a tobacco pathogen and in being lethal when heterologously expressed inside nonhost tobacco cells (Alfano et al., "Evidence That the *Pseudomonas syringae* pv. *syringae* *hrp*-linked *hrmA* Gene Encodes an Avr-like Protein that Acts in a *hrp*-dependent Manner Within Tobacco Cells," MPMI, 10:580-588 (1997), which is hereby incorporated by reference). The concept that the minimal requirement for bacterial elicitation of the HR is a functional Hrp system and an *avr* gene whose product is recognized by the *R*-gene surveillance system of the test plant is supported by experiments in which the HR is observed only when an appropriate, heterologous *avr* gene is supplied in trans of the *hrp*⁺ cosmid (Puri et al., "Expression of *avrPphB*, an Avirulence Gene from *Pseudomonas syringae* pv. *phaseolicola*, and the Delivery of Signals Causing the Hypersensitive Reaction in Bean," MPMI, 10:247-256 (1997); Ham et al., "A Cloned *Erwinia chrysanthemi* Hrp (type III Protein Secretion) System Functions in *Escherichia coli* to Deliver *Pseudomonas syringae* Avr Signals to Plant Cells and to Secrete Avr Proteins in Culture," Proc. Natl. Acad. Sci. USA, 95(17): 10206-11 (1998); Gopalan, S. et al., "Expression of the *Pseudomonas syringae* Avirulence Protein AvrB in Plant Cells Alleviates its Dependence on the Hypersensitive Response and pathogenicity (Hrp) Secretion System in Eliciting Genotype-specific Hypersensitive Cell Death," Plant Cell, 8:1095-1105 (1996); Pirhonen et al., "Phenotypic Expression of *Pseudomonas syringae* *avr* Genes in *E. Coli* is Linked to the Activities of the *hrp*-encoded Secretion System," MPMI 9:252-260 (1996), which are hereby incorporated by reference).

hrp genes are expressed in plants and in apoplast-mimicking minimal media, but typically not in complex media (Lindgren, "The Role of *hrp* Genes During Plant-Bacterial Interactions," Annu. Rev. Phytopathol., 35:129-152 (1997), which is hereby incorporated by reference). The Hrp regulatory systems in plant pathogenic bacteria can be divided into two groups, which correspond also to differences in *hrp* cluster composition (Alfano et al., "The Type III (Hrp) Secretion Pathway of Plant Pathogenic Bacteria: Trafficking Harpins, Avr Proteins, and Death," J. Bacteriol. 179:5655-5662 (1997), which is hereby incorporated by reference). In the group I Hrp systems of *Erwinia* and *Pseudomonas*, *hrp* operons are activated by HrpL, an alternate sigma factor (Lindgren, "The Role of *hrp* Genes During Plant-Bacterial Interactions," Annu. Rev. Phytopathol. 35:129-152 (1997); Hutcheson, "The *hrp*-encoded Protein Export Systems of *Pseudomonas syringae* and Other Plant Pathogenic Bacteria and Their

- Role in Pathogenicity," in Plant-Microbe Interactions, Volume 3, ed. Stacey and Keen, New York: Chapman and Hall, Inc., pp. 145-179 (1997), which are hereby incorporated by reference). In contrast, *hrp* transcription is activated by an AraC homolog in the group II Hrp systems of *Xanthomonas* (HrpX) and *Ralstonia* (HrpB) (Lindgren, "The
- 5 Role of *hrp* Genes During Plant-Bacterial Interactions," Annu. Rev. Phytopathol. 35:129-152 (1997), which is hereby incorporated by reference). Upstream activators of these factors have been described for *Pseudomonas syringae* (HrpR and HrpS, Φ^{54} -dependent promoter enhancer-binding protein homologs) (Lindgren, "The Role of *hrp* Genes During Plant-Bacterial Interactions," Annu. Rev. Phytopathol. 35:129-152 (1997);
- 10 Hutcheson, "The *hrp*-encoded Protein Export Systems of *Pseudomonas syringae* and Other Plant Pathogenic Bacteria and Their Role in Pathogenicity," in Plant-Microbe Interactions, Volume 3, ed. Stacey and Keen, New York: Chapman and Hall, Inc., pp. 145-179 (1997), which are hereby incorporated by reference), *Xanthomonas campestris* pv. *vesicatoria* (HrpG, OmpR homolog) (Wengelnik et al., "HrpG, a Key *hrp* Regulatory
- 15 Protein of *Xanthomonas campestris* pv. *vesicatoria* is Homologous to Two-component Response Regulators," MPMI 9:704-712 (1996), which is hereby incorporated by reference), and *R. solanacearum* (PrhA) (Marenda et al., "PrhA Controls a Novel Regulatory Pathway Required for the Specific Induction of *Ralstonia solanacearum hrp* Genes in the Presence of Plant Cells," Mol. Microbiol. 27:437-453 (1998), which is
- 20 hereby incorporated by reference). The recent discovery of PrhA is particularly significant because this homolog of TonB-dependent outer membrane siderophore receptors, which appears to act at the top of the Hrp regulatory hierarchy, is required specifically for induction of *hrp* genes in the presence of plant cells and for full virulence in Arabidopsis (Marenda et al., "PrhA Controls a Novel Regulatory Pathway Required
- 25 for the Specific Induction of *Ralstonia solanacearum hrp* Genes in the Presence of Plant Cells," Mol. Microbiol. 27:437-453 (1998), which is hereby incorporated by reference). In the host-promiscuous pathogen *Erwinia carotovora*, production of the *hrpN*-encoded harpin is activated by the quorum sensing signal, *N*-(3-oxohexanoyl)-L-homoserine lactone and negatively regulated by RsmA, two global regulators similarly controlling
- 30 exoenzyme production (Mukherjee et al., "Molecular Characterization and Expression of the *Erwinia carotovora hrpN_{Ecc}* gene, Which Encodes An Elicitor of the Hypersensitive Reaction," MPMI 10:462-471 (1997); Cui, Y. et al., "The RsmA⁻ Mutants of *Erwinia carotovora* subsp. *carotovora* Strain Ecc71 Overexpress *hrpN_{Ecc}* and Elicit a

Hypersensitive Reaction-like Response in Tobacco Leaves," MPMI, 9-6565-573 (1996), which are hereby incorporated by reference).

The protein or polypeptide must be compatible for secretion by the type III secretion system employed. By compatible, it is intended that the protein or polypeptide contain a secretion signal that can be recognized by the particular type III secretion system that is employed. The secretion signal enables the expressed protein or polypeptide to be recognized by the type III secretion system and transported via the expressed secretion system into the extracellular environment in which the recombinant host cells exist, i.e., culture medium.

Suitable secretion signals can be either an mRNA or a polypeptide fragment of a naturally-occurring protein secreted by the type III secretion system.

Compatible secretion signals can readily be determined for any particular type III secretion system that is to be employed. By identifying proteins that are normally secreted by the type III secretion system, it is possible to prepare deletion mutants missing various fragments of the full length protein that is normally secreted by the secretion system. Using labeled antibodies raised against epitopes of the various deletion fragments that are expressed (i.e., N-terminal epitopes, C-terminal epitopes, etc.), it is possible to identify deletion mutants that are secreted and those that are not secreted. Thus, protein domains necessary for secretion of the full length protein can be readily identified. Once the protein domains have been identified and sequenced, they can be utilized as secretion signals in fusion proteins of the present invention.

Typically, the secretion signal is an N-terminal domain of a protein that is normally secreted by the particular type III secretion system. An exemplary secretion signal is a 201 amino acid sequence from the N-terminal domain of the DspE protein of *Erwinia amylovora*, which has an amino acid sequence corresponding to SEQ. ID. No. 1 as follows:

	Met	Glu	Leu	Lys	Ser	Leu	Gly	Thr	Glu	His	Lys	Ala	Ala	Val	His	Thr	
	1				5					10					15		
30	Ala	Ala	His	Asn	Pro	Val	Gly	His	Gly	Val	Ala	Leu	Gln	Gln	Gly	Ser	
				20				25						30			
	Ser	Ser	Ser	Ser	Pro	Gln	Asn	Ala	Ala	Ala	Ser	Leu	Ala	Ala	Glu	Gly	
				35			40						45				
35	Lys	Asn	Arg	Gly	Lys	Met	Pro	Arg	Ile	His	Gln	Pro	Ser	Thr	Ala	Ala	
		50				55						60					

- 15 -

Asp Gly Ile Ser Ala Ala His Gln Gln Lys Lys Ser Phe Ser Leu Arg
 65 70 75 80
 5 Gly Cys Leu Gly Thr Lys Lys Phe Ser Arg Ser Ala Pro Gln Gly Gln
 85 90 95
 Pro Gly Thr Thr His Ser Lys Gly Ala Thr Leu Arg Asp Leu Leu Ala
 100 105 110
 10 Arg Asp Asp Gly Glu Thr Gln His Glu Ala Ala Ala Pro Asp Ala Ala
 115 120 125
 Arg Leu Thr Arg Ser Gly Gly Val Lys Arg Arg Asn Met Asp Asp Met
 130 135 140
 15 Ala Gly Arg Pro Met Val Lys Gly Gly Ser Gly Glu Asp Lys Val Pro
 145 150 155 160
 20 Thr Gln Gln Lys Arg His Gln Leu Asn Asn Phe Gly Gln Met Arg Gln
 165 170 175
 Thr Met Leu Ser Lys Met Ala His Pro Ala Ser Ala Asn Ala Gly Asp
 180 185 190
 25 Arg Leu Gln His Ser Pro Pro His Ile
 195 200

This amino acid sequence is encoded by a DNA molecule having a 603 base nucleic acid
 sequence from the gene or gene fragment coding for DspE. This DNA molecule has a
 30 nucleotide sequence corresponding to SEQ. ID. No. 2 as follows:

ATGGAATTAA AATCACTGGG AACTGAACAC AAGGCGGCAG TACACACAGC GCGCACAAC 60
 35 CCTGTGGGGC ATGGTGTTC CTTACAGCAG GGCAGCAGCA GCAGCAGCCC GCAAATGCC 120
 GCTGCATCAT TGGCGGCAGA AGGCAAAAAT CGTGGGAAAA TGCCGAGAAT TCACCAGCCA 180
 TCTACTGCGG CTGATGGTAT CAGCGCTGCT CACCAGCAAA AGAAATCCTT CAGTCTCAGG 240
 40 GGCTGTTTGG GGACGAAAAA ATTTTCCAGA TCGGCACCGC AGGGCCAGCC AGGTACCACC 300
 CACAGCAAAG GGGCAACATT GCGCGATCTG CTGGCGCGGG ACGACGGCGA AACGCAGCAT 360
 GAGGCGGCCG CGCCAGATGC GCGCGTTTG ACCCGTTCGG GCGGCGTCAA ACGCCGCAAT 420
 45 ATGGACGACA TGGCCGGGCG GCCAATGCTG AAAGGTGGCA GCGGCGAAGA TAAGGTACCA 480
 ACGCAGCAAA AACGGCATCA GCTGAACAAT TTTGGCCAGA TGCGCCAAAC GATGTTGAGC 540
 50 AAAATGGCTC ACCCGGCTTC AGCCAACGCC GCGATCGCC TGCAGCATTC ACCGCCGCAC 600
 ATC 603

The 201 amino acid secretion signal of *Erwinia amylovora* DspE is compatible with the harpin secretion system of *Erwinia amylovora*. Other secretion signals that are compatible with various type III secretion systems have been described in the art and other are continually being identified.

5 Human-pathogenic *Yersinia* spp. use the type III secretion system to deliver seven effector Yops into the host cytoplasm and another three Yops to the bacterial milieu (Hueck, "Type III Protein Secretion Systems in Bacterial Pathogens of Animals and Plants," Microbiol. Mol. Biol. Rev. 62:379-433 (1998), which is hereby incorporated by reference). All Yops carry an mRNA targeting signal in their first 15
10 codons (Anderson and Schneewind, "A mRNA Signal for the Type III Secretion of Yop Proteins by *Yersinia enterocolitica*," Science 278:1140-1143 (1997); Anderson and Schneewind, "Type III Machines of Gram-negative Pathogens: Injecting Virulence Factors into Host Cells and More," Curr. Opin. Microbiol. 2:18-24 (1999), which are hereby incorporated by reference). Fusion of the first 15 codons of YopE to an Npt
15 reporter is sufficient for type III secretion of the hybrid to the bacterial milieu, and mutations shifting the reading frame of these codons do not abolish secretion (Anderson and Schneewind, "A mRNA Signal for the Type III Secretion of Yop Proteins by *Yersinia enterocolitica*", Science 278:1140-1143 (1997), which is hereby incorporated by reference). Thus, the targeting information resides in the mRNA rather than the
20 encoded peptide. The mRNA targeting signal appears universal among effector proteins secreted by the type III pathways of animal pathogens and plant pathogens. For example, *Yersinia enterocolitica* strongly secretes AvrB and AvrPto, and *Escherichia coli*(pCPP2156) secretes (much less efficiently) YopE and YopQ (Anderson et al., "mRNA Signal Universal to Plant and Animal Pathogens," Proc. Natl. Acad. Sci. USA
25 (1999), which is hereby incorporated by reference). The first 15 codons of *avrB* and *avrPto* are necessary for secretion from *Escherichia coli*(pCPP2156), and they are sufficient to target an AvrPto₁₋₁₅-Npt hybrid for secretion in *Yersinia* and *Pseudomonas syringae* (Anderson et al., "mRNA Signal Universal to Plant and Animal Pathogens," Proc. Natl. Acad. Sci. USA (1999), which is hereby incorporated by reference). Frame-
30 shift mutations changing the peptide encoded by the AvrPto₁₋₁₅ mRNA do not prevent secretion of Npt in *Yersinia* (Anderson et al., "mRNA Signal Universal to Plant and Animal Pathogens," Proc. Natl. Acad. Sci. USA (1999), which is hereby incorporated by

reference). Thus, the mRNA signal recognized by type III secretion systems appears to be universal.

The protein or polypeptide can be a naturally secreted protein or polypeptide homologous to the type III secretion system (i.e., normally secreted by the source organism from which the type III secretion system was obtained) or heterologous to the type III secretion system (i.e., normally secreted by a source organism other than that from which the type III secretion system was obtained). By way of example, a naturally secreted protein or polypeptide homologous to the harpin secretion systems of *Erwinia amylovora* include, among others, DspE and HrpN. Exemplary naturally secreted proteins or polypeptides which are heterologous to the harpin secretion systems of *Erwinia* include the various *Pseudomonas syringae* Avr proteins.

Two classes of extracellular Hrp proteins have now been defined – harpins and pilins. Harpins are glycine-rich proteins that lack cysteine, are secreted in culture when the Hrp systems is expressed, and possess heat-stable HR elicitor activity when infiltrated into the leaves of tobacco and several other plants (Alfano et al., “Bacterial Pathogens in Plants: Life Up Against the Wall,” Plant Cell, 8:1683-1698 (1996), which is hereby incorporated by reference). Mutation of the prototypical *hrpN* harpin gene in *Erwinia amylovora* Ea321 strongly diminishes HR and pathogenicity phenotypes (Kim et al., “HrpW of *Erwinia amylovora*, a New Harpin That is a Member of a Proposed Class of Pectate Lyases,” J. Bacteriol. 180(19):5203-5210 (1998), which is hereby incorporated by reference), but mutation of the *hrpZ* harpin gene in different *Pseudomonas syringae* strains has little or no effect on Hrp phenotypes (Alfano et al., “Analysis of the Role of the *Pseudomonas syringae* pv. *syringae* HrpZ Harpin in Elicitation of the Hypersensitive Response in Tobacco Using Functionally Nonpolar Deletion Mutations, Truncated HrpZ Fragments, and *hrmA* Mutations,” Mol. Microbiol. 19:715-728 (1996); Charkowski et al., “The *Pseudomonas syringae* pv. *tomato* HrpW Protein Has Domains Similar to Harpins and Pectate Lyases and Can Elicit the Plant Hypersensitive Response and Bind to Pectate,” J. Bacteriol. 180 (19):5211-5217 (1998), which are hereby incorporated by reference). The natural function of harpins or the basis for their ability to elicit an apparent programmed cell death when artificially introduced into the apoplast of plants is unknown. However, two lines of evidence point to a site of action in the plant cell wall. First, purified *Pseudomonas syringae* harpin binds to cell walls and has biological activity only with walled cells (Hoyos et al., “The Interaction of

Harpin_{PS} With Plant Cell Walls," MPMI 9:608-616 (1996), which is hereby incorporated by reference). Second, HrpW, a second harpin discovered in both *Erwinia amylovora* and *Pseudomonas syringae*, has an N-terminal half that is harpin-like but a C-terminal half that is homologous to a newly-defined class of pectate lyases found in fungal and bacterial pathogens (Kim et al., "HrpW of *Erwinia amylovora*, a New Harpin That is a Member of a Proposed Class of Pectate Lyases," J. Bacteriol. 180(19):5203-5210 (1998); Charkowski et al., "The *Pseudomonas syringae* pv. *tomato* HrpW Protein Has Domains Similar to Harpins and Pectate Lyases and Can Elicit the Plant Hypersensitive Response and Bind to Pectate," J. Bacteriol. 180 (19):5211-5217 (1998), which are hereby incorporated by reference). Elicitor activity resides in the harpin domain, and the pectate lyase domain, although lacking enzymatic activity, binds specifically to pectate (Charkowski, A. et al., "The *Pseudomonas syringae* pv. *tomato* HrpW Protein Has Domains Similar to Harpins and Pectate Lyases and Can Elicit the Plant Hypersensitive Response and Bind to Pectate," J. Bacteriol. 180 (19):5211-5217 (1998), which is hereby incorporated by reference). The second class of extracellular Hrp proteins are represented by the *Pseudomonas syringae* HrpA pilin, which is a subunit of a Hrp-pilus that is 6-8 nm in diameter and is formed on bacteria in a Hrp-dependent manner (Roine et al., "Hrp Pilus: An *hrp*-dependent Bacterial Surface Appendage Produced by *Pseudomonas syringae* pv. *tomato* DC3000," Proc. Natl. Acad. Sci. USA 94:3459-3464 (1997), which is hereby incorporated by reference). The Hrp pilus is required for pathogenicity and elicitation of the HR, and a similar structure is important for T-DNA transfer in *Agrobacterium tumefaciens* (Fullner et al., "Pilus Assembly by *Agrobacterium* T-DNA Transfer Genes," Science, 237:1107-1109 (1996), which is hereby incorporated by reference). Whether these structures promote the transfer of bacterial macromolecules into plant cells by serving as conduits, guides, or attachment factors is not known.

A current model for plant-bacterium interaction and co-evolution based on Hrp delivery of Avr proteins into plant cells proposes that (i) Avr-like proteins are the primary effectors of parasitism, (ii) conserved Hrp systems are capable of delivering many, diverse Avr-like proteins into plant cells, and (iii) genetic changes in host populations that reduce the parasitic benefit of an effector protein or allow its recognition by the *R*-gene surveillance system will lead to a proliferation of complex arsenals of *avr*-like genes in co-evolving bacteria (Alfano et al., "Bacterial Pathogens in Plants: Life Up

Against the Wall," Plant Cell, 8:1683-1698 (1996), which is hereby incorporated by reference). There are still many gaps in this model. For example, the physical transfer of Avr proteins into plant cells has never been observed, the virulence functions of Avr proteins are unknown, and it is likely that previous searches for *avr* genes in various bacteria have yielded incomplete inventories of the genes in various bacteria and, thus, incomplete inventories of the genes encoding effector proteins.

Avr proteins have not been reported outside of the cytoplasm of living *Pseudomonas syringae* and *Xanthomonas* spp. cells (Leach et al., "Bacterial Avirulence Genes," Annul. Rev. Phytopathol., 34:153-179 (1996); Puri et al., "Expression of *avrPphB*, an Avirulence Gene from *Pseudomonas Syringae* pv. *phaseolicola*, and the Delivery of Signals Causing the Hypersensitive Reaction in Bean," MPMI 10:247-256 (1997), which are hereby incorporated by reference), but it now appears that the Hrp systems of *Erwinia* spp. can secrete Avr proteins in culture. A homolog of the *Pseudomonas syringae* pv. *tomato avrE* gene has been found in *Erwinia amylovora* and designated *dspA* in strain CFBP1430 and *dspE* in strain Ea321 (Gaudriault et al., "DspA, an Essential Pathogenicity Factor of *Erwinia amylovora* Showing Homology with AvrE of *Pseudomonas syringae*, is Secreted via the Hrp Secretion Pathway in a DspB-dependent Way," Mol. Microbiol., 26:1057-1069 (1997); Bogdanove et al., "Homology and Functional Similarity of a *hrp*-linked Pathogenicity Operon, *dspEF*, of *Erwinia amylovora* and the *avrE* locus of *Pseudomonas syringae* Pathovar Tomato," Proc. Natl. Acad. Sci. USA, 95:1325-1330 (1998), which are hereby incorporated by reference). *dsp* genes are required for the pathogenicity of *Erwinia amylovora*, but not for HR elicitation. A protein of the expected size of DspA is secreted in a Hrp- and DspB-dependent manner by CFBP1430 (DspB is a potential chaperone) (Gaudriault et al., "DspA, an Essential Pathogenicity Factor of *Erwinia amylovora* Showing Homology with AvrE of *Pseudomonas syringae*, is Secreted via the Hrp Secretion Pathway in a DspB-dependent Way," Mol. Microbiol., 26:1057-1069 (1997), which is hereby incorporated by reference). Specific antibodies were used to demonstrate unambiguously that DspE is efficiently secreted in a Hrp-dependent manner by strain Ea321 (Bogdanove et al., "*Erwinia amylovora* Secretes DspE, a Pathogenicity Factor and Functional AvrE Homolog, Through the Hrp (Type III Secretion) Pathway," J. Bacteriol., 180(8):2244-2247 (1998), which is hereby incorporated by reference).

Nothing is known of the localization or expected site of action of AvrE. However, there is strong evidence that the site of action of the *Pseudomonas syringae* AvrB and AvrPto proteins is inside plant cells (see Bonaset al., "Recognition of Bacterial Avirulence Proteins Occurs Inside the Plant Cell: A General Phenomenon in Resistance to Bacterial Diseases?," Plant J. 12:1 (1997); Baker et al., "Recognition and Signaling in Plant-Microbe Interactions," Science, 276:726-733 (1997), which are hereby incorporated by reference), and both proteins have now been found to be secreted by an *Erwinia chrysanthemi* Hrp system functioning heterologously in *Escherichia coli* (Ham et al., "A Cloned *Erwinia chrysanthemi* Hrp (type III Protein Secretion) System Functions in *Escherichia coli* to Deliver *Pseudomonas syringae* Avr Signals to Plant Cells and to Secrete Avr Proteins in Culture," Proc. Natl. Acad. Sci. USA, 95(17): 10206-11 (1998), which is hereby incorporated by reference). This secretion is Hrp-dependent, and *Escherichia coli* cells carrying the *Erwinia chrysanthemi* hrp genes also elicit an *avrB*-dependent HR in appropriate test plants. A strong implication of this work is that *Erwinia chrysanthemi*, which is a host-promiscuous soft-rot pathogen, also carries *avr*-like genes. The ability of the cloned *Erwinia chrysanthemi* Hrp system to secrete *Pseudomonas syringae* Avr proteins should promote searches for additional *avr*-like genes by providing a phenotype that is independent of plant tests, and it will enable direct investigation of Avr targeting signals and secretion mechanisms. For example, chaperone-independent targeting information in two *Yersinia* Yop proteins has been shown to reside in the mRNA encoding the N-terminus of the protein (Anderson et al., "A mRNA Signal for the Type III Secretion of Yop Proteins by *Yersinia Enterocolitica*," Science, 278:1140-1143 (1997), which is hereby incorporated by reference). The involvement of similar signals in Avr secretion is suggested by the need for continued protein (but not mRNA) synthesis *in planta* for Avr signal delivery, which would be consistent with a co-translational secretion process (Puri et al., "Expression of *avrPphB*, an Avirulence Gene from *Pseudomonas syringae* pv. *phaseolicola*, and the Delivery of Signals Causing the Hypersensitive Reaction in Bean," MPMI, 10:247-256 (1997), which is hereby incorporated by reference).

The biochemical activities or parasite-promoting functions of Avr proteins remain unclear, although several of those known make measurable contributions to virulence (Leach et al., "Bacterial Avirulence Genes," Annul. Rev. Phytopathol. 34:153-179 (1996), which are hereby incorporated by reference). Members of the

AvrBs3 family in *Xanthomonas* spp. are targeted to the plant nucleus (Van den Ackerveken et al., "Bacterial Avirulence Proteins as Triggers of Plant Defense Resistance," Trends Microbiol., (1997); Gabriel, "Targeting of Protein Signals from *Xanthomonas* to the Plant Nucleus," Trends Plant Sci., 2:204-206 (1997), which are
5 hereby incorporated by reference), and some of these have been shown recently to redundantly encode watersoaking functions associated with circulence (Yang et al., "Watersoaking Function(s) of XcmH1005 are Redundantly Encoded by Members of the *Xanthomonas avr/pth* Gene Family," MPMI, 9:105-113 (1996), which is hereby incorporated by reference. AvrD (*Pseudomonas syringae* pv. *tomato*) directs the
10 synthesis of syringolide elicitors of the HR (Leach et al., "Bacterial Avirulence Genes," Annul. Rev. Phytopathol., 34:153-179 (1996), which is hereby incorporated by reference); AvrBs2 (*Xanthomonas campestris* pv. *vesicatoria*) shows similarity to *A. tumefaciens* agrocinopine synthase (Swords et al., "Spontaneous and Induced Mutations in a Single Open Reading Frame Alters Both Virulence and Avirulence in *Xanthomonas*
15 *campestris* pv. *vesicatoria* *avrBs2*," J. Bacteriol., 4661-4669 (1996), which is hereby incorporated by reference); and AvrRxv (*Xanthomonas campestris* pv. *vesicatoria*) is a homolog of AvrA (*Salmonella typhimurium*) and YopJ (*Yersinia* spp.), proteins which travel the type III pathway in animal pathogens and trigger apoptosis in macrophages (Hardt et al., "A Secreted *Salmonella* Protein With Homology to an Avirulence
20 Determinant of Plant Pathogenic Bacteria," Proc. Natl. Acad. Sci. USA, 94:9887-9892 (1997); Monack et al., *Yersinia* Signals Macrophages to Undergo Apoptosis and YopJ is Necessary for this Cell Death," Proc. Natl. Acad. Sci. USA, 94:10385-10390 (1997), which are hereby incorporated by reference. This last observation has led to the suggestion that *avr-R* gene interactions may occur also in animal pathogenesis (Galan,
25 "Avirulence Genes' in Animal Pathogens?," Trends Microbiol., 6:3-6 (1998), which is hereby incorporated by reference.

The primary sequences of the *Pseudomonas syringae* Avr proteins reveal little about their potential function, but interestingly, when heterologously expressed in plants, three of them have produced necrosis in test plants lacking the cognate *R* gene
30 (Gopalan et al., "Expression of the *Pseudomonas syringae* Avirulence Protein AvrB in Plant Cells Alleviates its Dependence on the Hypersensitive Response and pathogenicity (Hrp) Secretion System in Eliciting Genotype-specific Hypersensitive Cell Death," Plant Cell, 8:1095-1105 (1996); Stevens et al., "Sequence Variations in Alleles of the

Avirulence Gene *avrPphE.R2* from *Pseudomonas syringae* pv. *phaseolicola* Lead to Loss of Recognition of the AvrPphE Protein Within Bean Cells and Gain in Cultivar Specific Virulence," Mol. Microbiol., 29(1):165-77 (1998); McNellis et al., "Glucocorticoid-inducible Expression of a Bacterial Avirulence Gene in Transgenic *Arabidopsis* Induces Hypersensitive Cell Death," Plant J., 14(2):247-57 (1998), which are hereby incorporated by reference). A key question is whether this results from interaction of abnormally high levels of the bacterial protein with plant virulence targets or with cross-reacting *R*-gene products. Further evidence suggesting that some *avr* genes in *Pseudomonas syringae* are beneficial to the bacteria in host plants is found in recent studies of *avrD* and *avrPphE*. Highly conserved, nonfunctional alleles of these genes have been retained in pathogens whose hosts would recognize the functional Avr product (Stevens et al., "Sequence Variations in Alleles of the Avirulence Gene *avrPphE.R2* from *Pseudomonas syringae* pv. *phaseolicola* Lead to Loss of Recognition of the AvrPphE Protein Within Bean Cells and Gain in Cultivar Specific Virulence," Mol. Microbiol., 29(1):165-77 (1998); Keith et al., "Comparison of *avrD* Alleles from *Pseudomonas syringae* pv. *glycinea*," MPMI, 10:416-422 (1997), which are hereby incorporated by reference).

Avr-like genes may function heterologously to support pathogenesis as well as HR elicitation. The pathogenicity of an *Erwinia amylovora dspE* mutant can be restored (at least partially) by a plasmid carrying the *Pseudomonas syringae avrE* locus, suggesting that DspE and AvrE have similar functions (Bogdanove et al., "Homology and Functional Similarity of a *hrp*-linked Pathogenicity Operon, *dspEF*, of *Erwinia amylovora* and the *avrE* locus of *Pseudomonas syringae* Pathovar Tomato," Proc. Natl. Acad. Sci. USA, 95:1325-1330 (1998), which is hereby incorporated by reference). That *dspE* is essential for *Erwinia amylovora* pathogenicity, whereas *avrE* contributes only quantitatively to the virulence of *Pseudomonas syringae* pv. *tomato* (Lorang et al., "*avrA* and *avrE* in *Pseudomonas Syringae* pv. *Tomato* PT23 Play a Role in Virulence on Tomato Plants," MPMI, 7:508-515 (1994), which is hereby incorporated by reference), suggests that there is less redundancy in the *Erwinia amylovora* virulence system. This would be consistent with a more recent acquisition of the Hrp system by *Erwinia amylovora* and/or a slower coevolution with its perennial hosts (Bogdanove et al., "Homology and Functional Similarity of a *hrp*-linked Pathogenicity Operon, *dspEF*, of *Erwinia amylovora* and the *avrE* locus of *Pseudomonas syringae* Pathovar Tomato,"

Proc. Natl. Acad. Sci. USA, 95:1325-1330 (1998), which is hereby incorporated by reference). The heterologous function of *Pseudomonas syringae* *avr* genes in *Erwinia amylovora* and *Erwinia chrysanthemi* suggests that Hrp+ bacteria in the field may be able to 'sample' a buffet of *avr*-like genes from diverse sources in their coevolution with changing plant populations. Many *avr* genes have been known to be potentially mobile, because of their presence on plasmids (Vivian et al., "Avirulence Genes in Plant-Pathogenic Bacteria: Signals or Weapons?," Microbiology 143:693-704 (1997); Leach et al., "Bacterial Avirulence Genes," Annul. Rev. Phytopathol., 34:153-179 (1996), which are hereby incorporated by reference). Recent observations with *Pseudomonas syringae* highlight the apparent mobility of *avr* genes. Several *Pseudomonas syringae* *avr* genes are linked with transposable elements or phage sequences (Hanekamp et al., "Avirulence Gene D of *Pseudomonas syringae* pv. *Tomato* May Have Undergone Horizontal Gene Transfer," FEBS Lett., 415:40-44 (1997), which is hereby incorporated by reference), and the *hrp* clusters in different strains of *Pseudomonas syringae*, although conserved in themselves, are bordered by a hypervariable region enriched in *avr* genes and mobile DNA elements.

Alternatively, the protein or polypeptide can be a protein or polypeptide which is not normally secreted by a type III secretion system, in which case the protein or polypeptide is expressed and secreted from the host cell as a fusion protein. The secretion signal, whether a mRNA or a polypeptide fragment, is linked to the protein or polypeptide of interest by an in-frame gene fusion, which preferably results in linking the mRNA or polypeptide fragment to the N-terminal end of the protein or polypeptide of interest. Such fusion proteins include a secretion signal linked to a protein or polypeptide of interest, for example, by a peptide bond between the secretion signal and the protein or polypeptide of interest. Fusion proteins can be prepared by ligating two or more DNA molecules together, one of which encodes the protein or polypeptide of interest and the other of which encodes the secretion signal. The two DNA molecules must be ligated in a manner which allows their proper expression. A number of efficient expression schemes for preparing fusion proteins have been developed and are well known in the art. According to one approach the fusion protein is prepared with a protease cleavage site intermediate the protein or polypeptide of interest and the secretion signal, such that the secretion signal can be removed from the protein or polypeptide of interest by, for example, proteolytic cleavage following isolation of the

fusion protein. A linker or spacer peptide may also be included to promote proteolytic cleavage (Polyak et al., "Introduction of Spacer Peptides N-terminal to a Cleavage Recognition Motif in Recombinant Fusion Proteins Can Improve Site-specific Cleavage," Protein Eng. 10(6):615-619 (1997), which is hereby incorporated by
5 reference). A number of suitable proteases (e.g., factor Xa, chymosin, trypsin, etc.) and their protease-specific cleavage sites are known in the art, and others continually being identified. Any protease capable of cleaving a specific amino acid sequence without disrupting the protein or polypeptide of interest can be employed.

Thus, DNA molecules encoding a protein or polypeptide capable of
10 being secreted by the type III secretion system can either be isolated (i.e., those encoding proteins that naturally contain a compatible secretion signal) or fabricated (i.e., those encoding for fusion proteins that contain a secretion signal linked by an in-frame gene fusion to a protein or polypeptide of interest). As indicated above, a single DNA construct of the present invention can contain both a type III secretion system and a
15 sequence encoding a protein or polypeptide capable of secretion by the type III secretion system or a pair of DNA constructs can be employed.

Regardless of which approach is utilized, once the DNA molecule is constructed and obtained, it can be incorporated in cells using conventional recombinant DNA technology. Generally, this involves inserting the DNA molecule into an
20 expression system to which the DNA molecule is heterologous (i.e., not normally present). Expression systems of the present invention contain an expression vector into which is inserted one or more heterologous DNA constructs of the present invention. The heterologous DNA molecule is inserted into the expression system or vector in proper sense orientation. The vector contains the necessary elements for the
25 transcription of the DNA constructs of the present invention.

U.S. Patent No. 4,237,224 to Cohen and Boyer, which is hereby incorporated by reference, describes the production of expression systems in the form of recombinant plasmids using restriction enzyme cleavage and ligation with DNA ligase. These recombinant plasmids are then introduced by means of transformation and
30 transfection, and replicated in cultures including prokaryotic organisms and eukaryotic cells grown in tissue culture.

Recombinant or engineered genes may also be introduced into viruses, such as vaccinia virus. Recombinant viruses can be generated by transfection of plasmids into cells infected with virus.

Suitable vectors include, but are not limited to, the following viral vectors
5 such as lambda vector system gt11, gt WES.tB, Charon 4, and plasmid vectors such as pBR322, pBR325, pACYC177, pACYC184, pUC8, pUC9, pUC18, pUC19, pLG339, pR290, pKC37, pKC101, SV 40, pBluescript II SK +/- or KS +/- (see "Stratagene Cloning Systems" Catalog (1993) from Stratagene, La Jolla, Calif, which is hereby incorporated by reference), pQE, pIH821, pGEX, pET series (see Studier et. al., "Use of
10 T7 RNA Polymerase to Direct Expression of Cloned Genes," Gene Expression Technology, vol. 185 (1990), which is hereby incorporated by reference), and any derivatives thereof. Suitable vectors are continually being developed and identified. Recombinant molecules can be introduced into cells via transformation, transduction, conjugation, mobilization, or electroporation. The DNA sequences are cloned into the
15 vector using standard cloning procedures in the art, as described by Maniatis et al., Molecular Cloning: A Laboratory Manual, Cold Springs Laboratory, Cold Springs Harbor, New York (1982), which is hereby incorporated by reference.

A variety of host-vector systems may be utilized to express the type III secretion system and the protein or polypeptide capable of secretion by the type III
20 secretion system. Primarily, the vector system must be compatible with the host cell used. Host-vector systems include, but are not limited to, the following: bacteria transformed with bacteriophage DNA, plasmid DNA, or cosmid DNA; microorganisms such as yeast containing yeast vectors; mammalian cell systems infected with virus (e.g., vaccinia virus, adenovirus, etc.); insect cell systems infected with virus (e.g.,
25 baculovirus); and plant cells infected by bacteria or transformed via particle bombardment (i.e., biolistics). The expression elements of these vectors vary in their strength and specificities. Depending upon the host-vector system utilized, any one of a number of suitable transcription elements can be used.

Different genetic signals and processing events control many levels of
30 gene expression (e.g., DNA transcription and messenger RNA (mRNA) translation).

Transcription of DNA is dependent upon the presence of a promoter which is a DNA sequence that directs the binding of RNA polymerase and thereby promotes mRNA synthesis. The DNA sequences of eukaryotic promoters differ from

those of prokaryotic promoters. Furthermore, eukaryotic promoters and accompanying genetic signals may not be recognized in or may not function in a prokaryotic system, and, further, prokaryotic promoters are not recognized and do not function in eukaryotic cells.

5 Specific initiation signals are also required for efficient gene transcription and translation in prokaryotic cells. These transcription and translation initiation signals may vary in "strength" as measured by the quantity of gene specific messenger RNA and protein synthesized, respectively. The DNA expression vector, which contains a promoter, may also contain any combination of various "strong" transcription and/or
10 translation initiation signals. Efficient translation of mRNA in prokaryotes requires a ribosome binding site called the Shine-Dalgarno ("SD") sequence on the mRNA. This sequence is a short nucleotide sequence of mRNA that is located before the start codon, usually ATG, which encodes the amino-terminal methionine of the protein. The SD sequences are complementary to the 3'-end of the 16S rRNA (ribosomal RNA) and
15 probably promote binding of mRNA to ribosomes by duplexing with the rRNA to allow correct positioning of the ribosome. Thus, any SD-ATG combination that can be utilized by host cell ribosomes may be employed. Such combinations include, but are not limited to, SD-ATG combinations synthesized by recombinant techniques, the SD-ATG combination from the *cro* gene or the N gene of coliphage lambda, or from the
20 *Escherichia coli* tryptophan E, D, C, B or A genes. For a review on maximizing gene expression, see Roberts and Lauer, *Methods in Enzymology*, 68:473 (1979), which is hereby incorporated by reference.

Promoters vary in their "strength" (i.e. their ability to promote transcription). For the purposes of expressing a cloned DNA construct of the present
25 invention, it is desirable to use strong promoters in order to obtain a high level of transcription and, hence, expression of the DNA construct. Depending upon the host cell system utilized, any one of a number of suitable promoters may be used. For instance, when cloning in *Escherichia coli*, its bacteriophages, or plasmids, promoters such as the T7 phage promoter, lac promoter, trp promoter, recA promoter, ribosomal RNA
30 promoter, the PR and PL promoters of coliphage lambda and others, including but not limited, to lacUV5, ompF, bla, lpp, and the like, may be used to direct high levels of transcription of adjacent DNA segments. Additionally, a hybrid trp-lacUV5 (tac) promoter or other *Escherichia coli* promoters produced by recombinant DNA or other

synthetic DNA techniques may be used to provide for transcription of the inserted construct.

Bacterial host cell strains and expression vectors may be chosen which inhibit the action of the promoter unless specifically induced. In certain operons, the addition of specific inducers is necessary for efficient transcription of the inserted DNA. For example, the *lac* operon is induced by the addition of lactose or IPTG (isopropylthio-beta-D-galactoside). A variety of other operons, such as *trp*, *pro*, etc., are under different controls.

Host cells can be transformed using the expression systems of the present invention, whereby the host cell is transformed with one or more of the DNA constructs of the present invention, as described above. Preferably, the host cells are present in a cell culture. Although any bacterial cell is suitable for use as a host cell, *Escherichia coli*, *Erwinia amylovora*, and *Erwinia chrysanthemi* are preferred host cells.

Biological markers can be used to identify the cells carrying recombinant DNA molecules. In bacteria, these are commonly drug-resistance genes. Drug resistance is used to select bacteria that have taken up cloned DNA from the much larger population of bacteria that have not. In the early mammalian gene transfer experiments involving viral genes, the transfer of exogenous DNA into cells was detected because the DNA had a biological activity, it led to production of infectious virus or produced stable changes in the growth properties of the transfected cells. It was then discovered that the DNA tumor virus, herpes simplex virus (HSV), contained a gene encoding the enzyme thymidine kinase (the *tk* gene). The HSV *tk* gene can be used as a selectable genetic marker in mammalian cells in much the same way that drug-resistance genes worked in bacteria, to allow rare transfected cells to grow up out of a much larger population that did not take up any DNA. The cells are transferred to selective growth medium, which permits growth only of cells that took up a functional *tk* gene (and the transferred DNA of interest). Various dominant selectable markers are now known in the art, including: aminoglycoside phosphotransferase (APH), using the drug G418 for selection which inhibits protein synthesis; the APH inactivates G418; dihydrofolate reductase (DHFR):Mtx-resistant variant, using the drug methotrexate (Mtx) for selection which inhibits DHFR; the variant DHFR is resistant to Mtx; hygromycin-B-phosphotransferase (HPH), using the drug hygromycin-B which inhibits protein synthesis; the HPH inactivates hygromycin B; thymidine kinase (TK), using the drug aminopterin which

inhibits de novo purine and thymidylate synthesis; the TK synthesizes thymidylate; xanthine-guanine phosphoribosyltransferase (XGPR1), using the drug mycophenolic acid which inhibits de novo GMP synthesis; XGPRT synthesizes GMP from xanthine; and adenosine deaminase (ADA), using the drug 9-b-D-xylofuranosyl adenine (Xyl-A) which damages DNA; the ADA inactivates Xyl-A. Other selectable markers are continually being identified.

Other aspects of the present invention relate to methods of secreting a protein or polypeptide into the environment of a host cell. According to one embodiment, this method is performed by introducing into a host cell a DNA construct of the present invention which contains both a nucleic acid sequence encoding a functional type III secretion system and a nucleic acid sequence encoding a protein or polypeptide capable of being secreted by the type III secretion system. The DNA construct is introduced into the host cell under conditions effective to cause expression of the encoded protein or polypeptide, wherein the encoded protein or polypeptide is secreted by the host cell into the environment (i.e., culture medium). According to a second embodiment, this method is performed by introducing into a host cell a two DNA construct system of the present invention, one of which contains a nucleic acid sequence encoding a functional type III secretion system and the other of which contains a nucleic acid sequence encoding a protein or polypeptide capable of being secreted by the type III secretion system. The DNA constructs are introduced into the host cell under conditions effective to cause expression of the encoded protein or polypeptide and the type III secretion system, wherein the encoded protein or polypeptide is secreted by the host cell into the environment.

Another aspect of the present invention relates to a method of isolating a protein or polypeptide of interest. This method is performed by first providing a recombinant host cell that contains (i) a DNA molecule encoding a functional type III secretion system and (ii) a heterologous DNA molecule having a promoter operably coupled to a nucleic acid sequence encoding a protein or polypeptide capable of being secreted by the type III secretion system. The recombinant host cell is then introduced into a culture medium, where the encoded protein or polypeptide and the type III secretion system are expressed and the encoded protein or polypeptide is secreted into the culture medium, and the encoded protein or polypeptide is isolated from the culture medium.

The recombinant host cell can contain a homologous type III secretion system, in which case the encoded protein or polypeptide to be secreted is heterologous to both the host cell and the type III secretion system. As indicated above, the encoded protein or polypeptide can be a naturally secreted protein or a fusion protein. For example, an *Erwinia amylovora* host cell (which contains a homologous hrp secretion system) can be transformed with a heterologous DNA molecule that contains a promoter operably coupled to a sequence encoding an Avr protein of *Pseudomonas syringae*. When the recombinant *Erwinia amylovora* is grown on appropriate culture medium, the Avr protein is expressed and secreted into culture medium, from which it is then isolated.

Alternatively, the recombinant host cell can contain a heterologous type III secretion system and a heterologous protein or polypeptide coding DNA molecule. The DNA constructs of the present invention can be used to transform a host cell. For example, an *Escherichia coli* host cell can be transformed with a single construct of the present invention which contains a DNA sequence encoding a functional type III secretion system of *Erwinia chrysanthemi* and a DNA sequence encoding an Avr protein of *Pseudomonas syringae*. When the recombinant *Escherichia coli* host cell is grown on appropriate culture media, the Avr protein is expressed and secreted into culture medium, from which it is then isolated.

Effective conditions include optimal growth temperatures and nutrient media which will enable maximal growth of the host cells and maximal expression of the protein or polypeptide of interest. Exemplary culture media include, without limitation, LM media and minimal media, both of which are known in the art. One of ordinary skill in the art can readily determine the optimal growth temperatures for particular strains of host cells and suitable nutrient media capable of optimizing host cell growth.

Purified protein may be obtained by several methods. The protein or polypeptide is preferably produced in purified form (preferably at least about 80%, more preferably 90%, pure) by conventional techniques. Since the recombinant host cells express a type III secretion system, the protein or polypeptide is secreted into the growth medium of recombinant host cells. In such cases, to isolate the protein, the recombinant host cells are propagated, the growth medium is centrifuged to separate cellular components from supernatant containing the secreted protein or polypeptide, and the supernatant is removed. The supernatant is then subjected to sequential ammonium sulfate precipitation. The fraction containing the polypeptide or protein is subjected to

gel filtration in an appropriately sized dextran or polyacrylamide column to separate the proteins. If necessary, the protein fraction may be further purified by HPLC.

Where the protein or polypeptide of interest is a fusion protein containing a protease-cleavable amino acid sequence between a secretion signal and a protein or polypeptide of interest, it is possible to isolate the fusion protein, expose the isolated fusion protein to a protease, and then recover (i.e., purify) the protein or polypeptide of interest as described above. Alternatively, to avoid repetitive purification steps, it is possible to introduce a protease into the culture medium or supernatant (i.e., following separation of the cellular component from the culture medium) prior to purification.

10 After treatment with the protease, the protein or polypeptide of interest can be purified as described above.

Because the constructs and systems of the present invention enable expression of a protein or polypeptide and its secretion into an environment (e.g., culture medium) of the host cell which contains the construct or system, the present invention

15 offers significant advantages in isolating a protein or polypeptide of interest. These advantages include the ability to dispense with disruption of the host cell membrane (i.e., by sonication or other known methods) to cause release of cellular contents into the culture medium and, thus, the ability to dispense with costly purification schemes for removal of cellular debris. By avoiding these practices, significant cost savings and

20 significant improvements in protein or polypeptide purity can be recognized.

Another aspect of the present invention relates to the ability to utilize the type III secretion system in a novel manner to prospect for putative effector proteins or polypeptides. As indicated previously, Avr proteins are one type of effector protein which have significant potential for use, for example, in the field of agriculture.

25 However, prior art methods of identifying potential effector proteins have proven to be inefficient.

Accordingly, this aspect of the present invention relates to a method of identifying a gene encoding a potential effector protein or polypeptide. This method is performed by first providing a host cell that contains a DNA molecule encoding a

30 functional type III secretion system. A gene to be screened (i.e., a candidate), which encodes a protein or polypeptide, is introduced into the host cell under conditions effective to express the encoded protein or polypeptide. Next, secretion of the encoded protein or polypeptide by the recombinant host cell is determined. Secretion of the

encoded protein or polypeptide (i.e., via the type III secretion system) indicates that the gene encodes a potential effector protein or polypeptide.

The recombinant host cell can contain a homologous type III secretion system, in which case the gene to be screened is heterologous to both the host cell and the type III secretion system. For example, an *Erwinia amylovora* host cell (which contains a homologous hrp secretion system) can be transformed with a heterologous gene obtained from *Pseudomonas syringae*.

Alternatively, the recombinant host cell can contain a heterologous type III secretion system and a heterologous gene to be screened. For example, an *Escherichia coli* host cell can be transformed with a DNA construct that contains a DNA sequence encoding a functional type III secretion system of *Erwinia chrysanthemi* and a DNA sequence encoding a gene obtained from *Pseudomonas syringae*. (Also, co-transformation of the host cell with two separate constructs can be performed.)

A preferred approach for determining whether the protein or polypeptide is expressed and secreted utilizes a chimeric gene that encodes an epitope tag fused to the protein or polypeptide. The gene to be screened can be a specific protein-coding gene or it can be obtained via shotgun cloning techniques. Regardless of how the gene is obtained, it is then modified, for example, according to the procedures of Gopalan et al., Plant Cell, 8:1095-1105 (1996), which is hereby incorporated by reference, to prepare the chimeric gene. The chimeric gene is prepared in a manner which preferably results in location of the epitope tag at the C-terminal end of the fusion protein. The recombinant host cell is grown in a suitable culture medium and then all protein secreted by the recombinant host cell is isolated and, preferably, immobilized. The isolated protein is then exposed to an immunodetection assay capable of recognizing the epitope tag.

Preferably the immunodetection assay utilizes a double antibody recognition complex, with the first antibody recognizing the epitope tag and the second antibody, which bears a detectable label, recognizing the first antibody. Together, the two antibodies enable detection of the epitope-tagged protein or polypeptide.

Monoclonal antibody production may be effected by techniques which are well-known in the art. A description of the theoretical basis and practical methodology for preparing hybridomas is set forth in Kohler and Milstein, *Nature* 256:495 (1975), which is hereby incorporated by reference. Procedures for raising polyclonal antibodies

are also well known, including procedures disclosed in E. Harlow, et. al., editors, Antibodies: A Laboratory Manual (1988), which is hereby incorporated by reference. In addition to utilizing whole antibodies, the processes of the present invention encompass use of binding portions of such antibodies. Such binding portions include Fab fragments, F(ab')₂ fragments, and Fv fragments. These antibody fragments can be made by conventional procedures, such as proteolytic fragmentation procedures, as described in Goding, Monoclonal Antibodies: Principles and Practice, pp. 98-118, New York:Academic Press (1983), which is hereby incorporated by reference.

Examples of labels useful for diagnostic imaging in accordance with the present invention are radioactive labels, fluorescent labels, nuclear magnetic resonance active labels, positron emitting isotopes detectable by a positron emission tomography ("PET") scanner, chemiluminescent labels, and enzymatic markers. A number of labels are well known in the art and others are continually being identified.

This method of screening for putative effector proteins enables one of ordinary skill in the art to more readily identify putative effector proteins, which can then be tested on host organisms to determine the ability of the putative effector protein to induce a host plant response.

These aspects of the present invention are further illustrated by the examples below.

EXAMPLES

The following examples are provided to illustrate embodiments of the present invention but they are by no means intended to limit its scope.

Materials and Methods

Bacterial strains, culture conditions, and DNA manipulation techniques:

Bacterial strains and plasmids used in this study are listed in Table 1 below.

Table 1: Bacterial Strains and Plasmids

Designation	Relevant Characteristics and Use	Reference or Source
<i>Escherichia coli</i> DH5α	SupE44 ΔlacU169 (f80 lacZΔM15) hsdR17 recA1 endA1 gyrA96 thi-1	Life Technologies

Designation	Relevant Characteristics and Use	Reference or Source
	relA1, N ^x	
XL0LR	$\Delta(mcrA)183 \Delta(mcrBc-hsdSMR-mrr)173 \text{ endA1 } thi-1 \text{ recA1 } gyrA96 \text{ relA1 } lac [F' \text{ proAB } laqI^qZ\Delta M15 \text{ Tn10}(Tc^r)]$	Stratagene
SM10 λ pir	SM10 lysogenized with λ -pir for mobilizing pUT::mini-Tn5Cm, Km ^r	(Miller et al., <u>J. Bacteriol.</u> , 170: 2575-2583 (1988), which is hereby incorporated by reference)
MC4100	F' araD139 $\Delta(argF-lacZYA)U169 \text{ rpsL } 150 \text{ relA1 } fib-5301 \text{ ptsF25 } deoC1$	(Casadaban, <u>J. Mol. Biol.</u> 104:541-555 (1976), which is hereby incorporated by reference)
Plasmids		
pFLAG-CTC	For construction of C-terminal fusion to FLAG peptide, Ap ^r	Kodak Scientific Imaging Systems
pUT::mini-Tn5Cm	Mini-Tn5 transposon with Cm ^r on suicide plasmid pGP704 derivative for transposon mutagenesis, Ap ^r	(de Lorenzo et al., <u>J. Bacteriol.</u> , 172:6568-6572 (1990), which is hereby incorporated by reference)
pML123	Broad host range expression vector for cloning <i>avrPto</i> -FLAG, Gm ^r	(Labes et al., <u>Gene</u> , 89:37-46 (1990), which is hereby incorporated by reference)
pPtE6	An <i>avrPto</i> clone in pDSK519, Km ^r	(Ronald et al., <u>J. Bacteriol.</u> 174: 1604-1611 (1992), which is hereby incorporated by reference)
pHIR11	pLAFR3 carrying Pss61 <i>hrp/hrc</i> cluster, Tc ^r	(Huang et al., <u>J. Bacteriol.</u> , 170: 4748-4756 (1988), which is hereby incorporated by reference)
pCPP2156	pCPP19 carrying <i>E. chrysanthemi hrp</i> cluster, Sp ^r	This work
pCPP2157	pCPP19 carrying <i>E. chrysanthemi hrp</i> cluster, Sp ^r	(Bauer et al., <u>MPMI</u> 8:484-491 (1995), which is hereby incorporated by reference)
pCPP2329	pFLAG-CTC carrying <i>avrPto</i> , Ap ^r	This work
pAVRB-FLAG2	pML 123 carrying <i>avrB</i> -FLAG, Gm ^r	(Gopalan et al., <u>Plant Cell</u> , 8:1095-1105 (1996), which is hereby incorporated by reference)
pAVRB1	pDSK519 carrying <i>avrB</i> , Km ^r	(Tamaki et al., <u>J. Bacteriol.</u> 170:4846-4854 (1988), which is hereby incorporated by reference) N.T. Keen
pAVRPTO-FLAG	pML 123 carrying <i>avrPto</i> -FLAG, Gm ^r	This work
pCPP2368	A pCPP2156::Tn5Cm that has HR ^r	This work

Designation	Relevant Characteristics and Use	Reference or Source
	phenotype, Sp ^r , Cm ^r	
pCPP2416	A pCPP2156::Tn5Cm that has HR ⁺ phenotype, Sp ^r , Cm ^r	This work
pCPP2318	pCPP30 carrying mature blaM, Tc ^r	(Charkowski et al., <u>J. Bacteriol.</u> 179:3866-3874 (1997), which is hereby incorporated by reference)

Escherichia coli strains were routinely grown in LM medium (Hanahan, D., J. Mol. Biol., 166:557-580 (1983), which is hereby incorporated by reference) at 37°C for isolation of plasmids and at 30°C for protein secretion assays. The following

- 5 concentrations of antibiotics were used in selective media: ampicillin (Ap), 100 µg/ml; chloramphenicol (Cm), 20 µg/ml; gentamycin (Gm), 10 µg/ml; kanamycin (Km), 50 µg/ml; nalidixic acid (Nx), 20 µg/ml; spectinomycin (Sp), 50 µg/ml; tetracycline (Tc), 10 µg/ml. Standard procedures were followed by DNA manipulations (Sambrook, J., et al., Molecular Cloning: A Laboratory Manual, 2nd Edition (Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY), pp Pages (1989), which is hereby incorporated by reference).

Mini-Tn5Cm mutagenesis of the *hrp* gene cluster in pCPP2156:

- Mini-Tn5Cm mutagenesis of *Escherichia coli* DH5α(pCPP2156) was initiated by conjugation with *Escherichia coli* SM10λpir (pUT::mini-Tn5Cm). Because
- 15 pUT cannot replicate in *Escherichia coli* DH5α, Cm^r transconjugants have mini-Tn5Cm transposed to the chromosome or pCPP2156. To obtain a pool of pCPP2156::mini-Tn5Cm plasmids, all Cm^r colonies were triparentally mated with *Escherichia coli* XL0LR (Tc^r). The cosmids from *Escherichia coli* XL0LR transconjugants, selected on LM agar containing Tc, Sp, and Cm, were isolated and their restriction fragment patterns
- 20 compared with each other and pCPP2156. All 46 cosmids initially examined contained random insertions of mini-Tn5Cm in pCPP2156. Two primers were used to sequence from both ends of mini-Tn5Cm including a first primer having a nucleotide sequence corresponding to SEQ. ID. No. 3 as follows:

AGATCTGATC AAGAGACAG

19

- 25 and a second primer having a nucleotide sequence corresponding to SEQ. ID. No. 4 as follows:

CCGTGTGTAT AAGAGTCAG

19

Based on restriction mapping and DNA sequencing from both ends of mini-Tn5Cm, two different pCPP2156:mini-Tn5Cm derivatives were chosen. In one of them, Tn5Cm was inserted in the intergenic region between *hrpJ* and *hrcV* in the *hrpJ* operon. This cosmid was named pCPP2368. The other cosmid contained mini-Tn5Cm outside of the *hrp* cluster and was named pCPP2416. Both cosmids were electroporated into *Escherichia coli* DH5 α .

Plant bioassays:

Tobacco (*Nicotiana tabacum* L. cv. Xanthi) and *Nicotiana clevelandii* were grown under greenhouse conditions and then maintained in the lab at room temperature with daylight and supplemental metal halide illumination for HR assays. Soybean (*Glycine max* L.) and tomato (*Lycopersicum esculentum* Mill. cv. Rio Grande) plants were grown from seeds in pots with Cornell Mix (Cornell University) in the lab at room temperature. *Escherichia coli* DH5 α cells grown overnight on LM plates were washed twice with 5mM MES (morpholinoethanesulfonic acid, pH 6.5) by centrifugation and then resuspended in an appropriate volume of the same buffer to an OD₆₀₀ of 0.8 (experiments involving *avrPto* used *Escherichia coli* MC4100 and a 3-fold higher level of inoculum). Previously described procedures were used for the infiltration of bacterial cells into tobacco, tomato and *Nicotiana clevelandii* leaves (Bauer et al., MPMI, 7:573-581 (1994), which is hereby incorporated by reference) and soybean leaves (Gopalan et al., Plant Cell, 8:1095-1105 (1996), which is hereby incorporated by reference).

Preparation of AvrB antibodies:

AvrB-FLAG was purified from *Escherichia coli* DH5 α (pFLAG-CTC::AvrB) by affinity chromatography as described (Gopalan et al., Plant Cell, 8:1095-1105 (1996), which is hereby incorporated by reference), followed by precipitation of aliquots containing 1 mg of partially purified protein with trichloroacetic acid ("TCA") (20% final concentration), resuspension in SDS polyacrylamide gel loading buffer (Sambrook, J., et al., Molecular Cloning: A Laboratory Manual, 2nd Edition (Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY), (1989), which is hereby incorporated by reference), and electrophoresis on 1.5 mm X 11 cm X 10 cm 12% polyacrylamide preparative gels. The AvrB band was excised from each gel following brief staining with a solution of 0.2% Coomassie R350 (Pharmacia Biotech) dissolved in water. Subsequent extraction of AvrB from the gel matrix and generation of polyclonal

rabbit anti-AvrB antisera were performed by the Immunological Resource Center at the University of Illinois, Urbana IL. Prior to usage, the antisera was delipified with sodium dextran sulfate (average molecular weight of 500,000) to a final concentration of 0.25% and CaCl₂ to a final concentration of 1.0% followed by incubation of 4°C for 8-12 hours (Walton, K.W., et al., J. Clin. Pathol., 17:627-643 (1964), which is hereby incorporated by reference). This mixture was clarified by centrifugation at 12,000 X g at 4°C for 10 minutes. Proteins were precipitated by the addition of ammonium sulfate at 50% saturation, followed by incubation at 4°C for 8-12 hours, and then collected by centrifugation at 12,000 X g at 4°C for 10 minutes and resuspended in their original volume with phosphate-buffered saline.

Construction of pAVRPTO-FLAG:

The *avrPto* gene was isolated by polymerase chain reaction ("PCR") with *Pfu* DNA polymerase (Stratagene, La Jolla, CA) and pPtE6 as the template. The upper primer had a nucleotide sequence that contains an *Nde* I site, corresponding to SEQ. ID. No. 5 as follows:

GAGCGAGCAT ATGGGAAATA TATGTGTCGG C 31

The lower primer had a nucleotide sequence that contains an *Sal* I site, corresponding to SEQ. ID. No. 6 as follows:

ATTGTAGTCG ACTTGCCAGT TACGGTACGG G 31

The reaction products from 30 PCR cycles were resolved by electrophoresis through 0.7% agarose, and the *avrPto* DNA was isolated using an Eluquick kit (Schleicher & Schuell), followed by digestion with *Nde* I and *Sal* I. This DNA was cloned into pFLAG-CTC, previously digested with *Nde* I and *Sal* I, and named pCPP2329. The *avrPto-FLAG* DNA was isolated from pCPP2329 by digestion with *Ssp* I and cloned into pML 123, which had been previously digested with *Bam*HI and blunted with Klenow polymerase, producing pAVRPTO-FLAG. As with *avrB-Flag2* (Gopalan, S., et al., Plant Cell, 8:1095-1105 (1996), which is hereby incorporated by reference), *avrPto-FLAG* is expressed by both the tax promoter (from pFLAG-CTC) and the pML 123 *nptII* promoter, thus permitting constitutive expression in LM medium and *in planta*.

Preparation of protein samples from supernatant and cell fractions:

Bacteria grown overnight on LM plates at 37°C were washed twice by centrifugation and resuspended in LM broth. Each bacterial suspension was diluted to

OD₆₀₀ = 0.2 in 40 ml of LM broth containing appropriate antibiotics and cultured at 30°C in a rotary shaking incubator at 220 rpm until the OD₆₀₀ reached 0.8. Centrifugations for the separation of bacterial cultures into cell-bound and supernatant fractions were performed with an SS-34 rotor (DuPont Instrument) at 4°C.

- 5 Forty milliliters of culture was initially centrifuged at 6,000 rpm (> 4300 x g) for 15 minutes. For the supernatant fraction, the upper 20 ml of supernatant was carefully transferred to a new centrifuge tube and further centrifuged at 12,000 rpm (> 17200 x g) for 40 minutes, followed by transfer of the upper 10 ml of supernatant to a new tube. Six milliliters of 25% TCA was added to the supernatant fraction, which was
- 10 then kept on ice for 3-4 hours, followed by centrifugation at 12,000 (> 17200 x g) rpm for 40 minutes. The pellet was subsequently washed with 20 ml of ice-cold acetone and then resuspended in 200 µl or 100 µl X SDS sample buffer (New England Biolabs). For the cell fraction, the pellet from the initial centrifugation was resuspended in 4 ml of LM broth. One hundred microliters of bacterial cell suspension was mixed with 50 µl of 3 X
- 15 SDS sample buffer. Each protein sample was held in a boiling water bath for 5 minutes before electrophoresis, and then 15 µl of each sample was loaded onto the gel.

Immunoblot analysis:

- Protein samples were separated by electrophoresis through a 10% SDS-polyacrylamide gel. Proteins in the gel were then electrotransferred to Immobilon-P
- 20 membrane (Millipore Corp.) with a Semi-Phor system (Hoefer Scientific Instruments). AvrB-FLAG and AvrPto-FLAG were detected with the Western-Light Plus kit (Tropix) using anti-FLAG M2 antibodies (Kodak Scientific Imaging Systems) and anti-mouse IgG alkaline phosphatase conjugate (Sigma) as primary and secondary antibodies, respectively. AvrB and β-lactamase were detected with the same system except using
- 25 anti-AvrB antibodies or anti-β-lactamase antibodies (5 prime → 3 prime) and anti-rabbit IgG alkaline phosphatase conjugate (Sigma) as primary and secondary antibodies, respectively.

Primers, DNA sequencing and data analysis:

- Oligonucleotide synthesis and DNA sequencing were performed at the
- 30 Cornell Biotechnology Center. DNA sequence data were managed and analyzed with the DNASTar Program (DNASTar, Madison, WI).

Example 1 - Preparation of Cosmid pCPP2156, Which Carries an Intact *Erwinia chrysanthemi* *hrp* Gene Cluster

Several cosmids carrying *Erwinia chrysanthemi* EC16 *hrp* genes were previously isolated on the basis of their ability to hybridize with an *Erwinia amylovora* DNA fragment carrying the *hrpJ* operon (Bauer, D.W., et al., MPMI, 7:573-581 (1994), which is hereby incorporated by reference). One of these cosmids, pCPP2157, is shown in Figure 1 and was subsequently found to carry also *hrpN* and the complete *hrpC* operon (Bauer, et al., "Erwinia chrysanthemi harpin_{Ech}: An Elicitor of the Hypersensitive Response that Contributes to Soft-rot Pathogenesis," 8:484-491 (1995); Kim, et al., "The *hrpC* and *HrpN* Operons of *Erwinia chrysanthemi* EC16 are Flanked by *plcA* and Homologs of Hemolysin/Adhesion Genes and Accompanying Activator/Transporter Genes," MPMI, 11(6):563-567 (1998), which are hereby incorporated by reference). Although pCPP2157 appeared to carry both borders of the *Erwinia chrysanthemi* *hrp* cluster, *Escherichia coli*(pCPP2157) failed to elicit an HR in tobacco leaves. DNA sequencing of the right end of the pCPP2157 insert revealed that *hrcU* was missing the last 180 nucleotides, as suggested by comparison with the *Erwinia amylovora* *hrcU* gene (Bogdanove, et al., "Erwinia amylovora Secretes Harpin via a Type III Pathway and Contains a Homolog of YopN of Yersinia spp.," J. Bacteriol., 178:1720-1730 (1996), which is hereby incorporated by reference). *hrcU* is one of nine former *hrp* genes that encode core components of the type III secretion system, are broadly conserved in plant and animal pathogenic bacteria, and have been renamed as *hrc* (HR and conserved) genes (Bogdanove, et al., Mol. Microbiol., 20:681-683 (1996), which is hereby incorporated by reference). Because of the *hrcU* truncation, additional cosmids hybridizing with probes carrying *hrpN* and *hrcU* were analyzed. pCPP2156 was one of those. Partial DNA sequence analysis and physical map comparisons with the *Erwinia amylovora* *hrp* genes suggested that pCPP2156 carried the entire *Erwinia chrysanthemi* *hrp* gene cluster, including at least one intercalated region not obviously related to Hrp function and 14-kb of additional DNA beyond *hrcU* (Figure 1). However, pCPP2156 failed to elicit an HR in tobacco.

Example 2 - *Escherichia coli*(pCPP2156) Enables Elicitation of an AvrB-dependent HR in *Nicotiana clevelandii* and Soybean Cultivar Norchief

The plasmid pHIR11, which carries the intact *Pseudomonas syringae* pv *syringae* 61 *hrp* cluster, enables *Escherichia coli* to elicit an HR in tobacco, because it also carries *hrmA*, an *avr*-like gene whose transient expression in tobacco cells is lethal (Alfano, et al., Mol. Microbiol., 19:715-728 (1996); Alfano, et al., MPMI, 10:580-588 (1997), which are hereby incorporated by reference). This suggested the possibility that pCPP2156 failed to elicit an HR in tobacco because it did not carry an appropriate *avr* gene. To test this, pAVRB-FLAG2 was transformed into *Escherichia coli* DH5 α cells carrying either pCPP2156 or pCPP2157. pAVRB-FLAG2 expresses the *Pseudomonas syringae* pv *glycinea* *avrB* gene such that the product has an eight-amino acid FLAG epitope C-terminal fusion (Gopalan, et al., Plant Cell, 8:1095-1105 (1996), which is hereby incorporated by reference). Transformants were infiltrated at a concentration of 5 X 10⁸ cfu/ml into *Nicotiana clevelandii*, a plant that reacts hypersensitively to Hrp⁺ bacteria carrying *avrB*. A typical HR developed within 24 hours in panels inoculated with bacteria carrying both *avrB* and pCPP2156, but there was no response in panels inoculated with bacteria lacking *avrB* or carrying an incomplete *hrp* cluster (Figure 2A), or inoculated with bacteria carrying only pAVRB-FLAG2.

The failure of *Escherichia coli*(pCPP2157, pAVRB-FLAG2) to elicit an HR in *Nicotiana clevelandii* suggested that this ability was Hrp-dependent. However, an explanation based on differences in the DNA flanking the *hrp* gene clusters in pCPP2156 and pCPP2157 remained a formal possibility. To resolve this, pCPP2156 was mutated with mini-Tn5Cm and two derivatives were isolated. Restriction mapping and DNA sequence analysis revealed that pCPP2416 and pCPP2368 carried insertions in the 14-kb region beyond *hrcU* and in the intergenic region between *hrpJ* and *hrcV*, respectively (Figure 1). The mutation in pCPP2368 would be expected to block transcription of *hrcV* and downstream genes in the putative *hrpJ* operon, and a polar mutation in the *hrpJ* operon of *Pseudomonas syringae* pv *syringae* 61 has been shown to result in accumulation of the HrpZ harpin within the bacterial cytoplasm (Charkowski et al., "Altered Localization of HrpZ in *Pseudomonas syringae* pv. *syringae* *hrp* Mutants Suggests That Different Components of the Type III Secretion Pathway Control Protein Translocation Across the Inner and Outer Membranes of Gram-negative Bacteria," J. Bacteriol., 179:3866-3874 (1997), which is hereby incorporated by reference).

Escherichia coli cells carrying pCPP2416 and pCPP2368 were transformed with pAVRB-FLAG2 and tested for their ability to elicit the HR in *Nicotiana clevelandii*. An HR was observed only with *Escherichia coli* (pCPP2416, pAVRB-FLAG2) (Figure 2A).

To further test the ability of *Escherichia coli* (pCPP2156) to elicit an
5 *avrB*-dependent HR, the leaves of soybean cultivars Norchief (*RPG1*) and Acme (*rpg1*) were infiltrated with bacteria carrying various plasmids. *RPG1* is an *R* gene that interacts in a gene-for-gene manner with *avrB* (Staskawicz, et al., *J. Bacteriol.*, 169:5789-5794 (1987), which is hereby incorporated by reference). The HR was observed only in Norchief inoculated with bacteria carrying both *avrB* and an intact *hrp*
10 cluster (Figure 2B).

Example 3 - *Escherichia coli*(pCPP2156) Secretes AvrB in Culture in a Hrp-dependent Manner while Retaining β -lactamase

The secretion of Avr proteins by *Pseudomonas syringae* is presumed to be
15 dependent on host cell contact, because Avr proteins remain cytoplasmic in culture even when the Hrp system is actively secreting harpins (Alfano, et al., *J. Bacteriol.*, 179:5655-5662 (1997), which is hereby incorporated by reference). The seemingly less specialized interaction of *Erwinia chrysanthemi* with its hosts suggested the possibility that Avr secretion may be regulated less tightly. To test this, FLAG epitope-tagged AvrB
20 encoded by pAVRB-FLAG2 was used to determine if *Escherichia coli* (pCPP2156) could secrete AvrB in culture. Bacterial cultures in late logarithmic-phase growth were separated into supernatant and cell-bound fractions by centrifugation, and then proteins in both fractions were resolved by SDS-PAGE. AvrB-FLAG was visualized by immunoblotting with anti-FLAG monoclonal antibodies and chemiluminescent
25 detection. AvrB-FLAG was found in the supernatant of *Escherichia coli*(pCPP2156) (Figure 3). Although much of the AvrB remained in the cell-bound fraction, secretion was Hrp-dependent and specific in that no AvrB-FLAG was found in the supernatant of *Escherichia coli* (pCPP2368) (Figure 3), and Coomassie staining revealed equally low levels of protein in the supernatant fractions of all of the bacteria tested.

30 To confirm that the presence of AvrB-FLAG in the *Escherichia coli*(pCPP2156) medium resulted from specific secretion and not cell lysis, and that secretion was not due to FLAG epitope, the localization of AvrB and mature β -lactamase was simultaneously monitored (Figure 4). *Escherichia coli* cells carrying pCPP2156 or pCPP2368 were first transformed with pAVRb1, which expresses AvrB from the *lac*

promoter (Tamaki, S., et al., J. Bacteriol., 170:4846-4854 (1988), which is hereby incorporated by reference), and pCPP2318, which encodes a mature β -lactamase that lacks its N-terminal signal peptide and can be used as a cytoplasmic marker (Charkowski et al., "Altered Localization of HrpZ in *Pseudomonas syringae* pv. *syringae* hrp Mutants Suggests That Different Components of the Type III Secretion Pathway Control Protein Translocation Across the Inner and Outer Membranes of Gram-negative Bacteria," J. Bacteriol., 179:3866-3874 (1997), which is hereby incorporated by reference). The distribution of AvrB and β -lactamase in the same supernatant and cell-bound fraction samples was monitored by immunoblotting with appropriate antibodies. The *Escherichia coli*(pCPP2156) supernatant sample contained AvrB but no β -lactamase (Figure 4), indicating that AvrB secretion occurred without the FLAG epitope and without cell lysis.

Example 4 - *Escherichia coli*(pCPP2156) Secretes AvrPto in Culture in a Hrp-dependent Manner

The evidence for Avr action inside plant cells following Hrp-dependent transfer is strongest with AvrPto, whose structural gene was originally isolated from *Pseudomonas syringae* pv. *tomato* (Alfano, et al., J. Bacteriol., 179:5655-5662 (1997); Tang, et al., Science, 274:2060-2062 (1996); Scofield, et al., Science, 274:2063-2065 (1996); Ronald, et al., J. Bacteriol., 174:1604-1611 (1992), which are hereby incorporated by reference). Consequently, AvrPto was selected to test whether the ability of *Escherichia coli*(pCPP2156) to deliver *Pseudomonas syringae* Avr signals *in planta* and to secrete Avr proteins in culture would extend beyond AvrB. First, the construct pAVRPTO-FLAG, which encodes AvrPto with a C-terminal FLAG epitope fusion, was prepared. *Escherichia coli* cells carrying pAVRPTO-FLAG and pCPP2156 (but not pAVRPTO-FLAG alone, pCPP2156 alone, or pAVRPTO-FLAG with pCPP2368) elicited an HR in tomato cultivar Rio Grande carrying the *Pto* resistance gene. The secretion of AvrPto was determined with the same methods used for AvrB-FLAG and AvrB. AvrPto-FLAG was secreted by *Escherichia coli*(pCPP2156) but not by *Escherichia coli*(pCPP2157) or *Escherichia coli*(pCPP2368) (Figure 5). Thus, secretion of AvrPto-FLAG was Hrp-dependent, and it also occurred without leakage of β -lactamase. In contrast, no AvrPto-FLAG was found in the supernatant of *Escherichia coli*(pHIR11), which expresses the intact *Pseudomonas syringae* Hrp system (Figure 5).

The isolation of a cluster of *Erwinia chrysanthemi* *hrp* genes that directs *Escherichia coli* to secrete *Pseudomonas syringae* Avr proteins in culture and deliver Avr signals in planta has several implications for the pathogenic biology of *Erwinia chrysanthemi* and *Pseudomonas syringae*. As representative necrotrophic and biotrophic parasites, respectively, these two bacteria mark the extremes in the pathogenic personalities of the common gram-negative phytopathogenic bacteria. Nevertheless, it is possible that they elicit the HR and initiate parasitic attack in fundamentally similar ways, that they may be able to interchange their *avr* genes without loss of function, and that their cloned *hrp* clusters can be used to biochemically investigate Avr protein secretion and to systematically prospect for the proteins injected into plants by many plant pathogenic bacteria.

Erwinia chrysanthemi and *Pseudomonas syringae* appear to elicit the HR by the same mechanism in that their cloned *hrp* clusters are dependent on an appropriate *avr* gene for elicitation of the HR when heterologously expressed in nonpathogens. Thus, cosmid pHIR11 (*Pseudomonas syringae* *hrp* cluster) directs HR elicitation in tobacco, because it carries the *avr*-like *hrmA* gene. Cosmid pCPP2156 (*Erwinia chrysanthemi* *hrp* cluster) fails to elicit the Hr in tobacco, *Nicotiana clevelandii*, soybean, or tomato because it does not carry an *avr* gene that is recognized by these plants, but when provided with *avrB* or *avrPto* it appropriately directs elicitation of the HR in *Nicotiana clevelandii*, soybean cultivar Norchief, and tomato cultivar Rio Grande. This has two implications regarding HR elicitation by *Erwinia chrysanthemi*. First, the harpin encoded by pCPP2156, like that encoded by pHIR11, is apparently insufficient for bacterial HR elicitation (although both harpins can elicit programmed cell death when delivered exogenously) (Bauer, et al., "Erwinia chrysanthemi harpin_{Ech}: An Elicitor of the Hypersensitive Response that Contributes to Soft-rot Pathogenesis," 8:484-491 (1995); He, et al., *Cell*, 73:1255-1266 (1993), which are hereby incorporated by reference). Second, *Erwinia chrysanthemi* must carry *avr* genes somewhere outside the region cloned in pCPP2156 because it is able to elicit a Hrp-dependent HR in tobacco without provision of a heterologous *avr* gene (Bauer, et al., *MPMI*, 7:573-581 (1994), which is hereby incorporated by reference).

Erwinia chrysanthemi *hrp* mutants also are reduced in their ability to elicit infection at low levels of inoculum. Because it now appears that the primary function of the Hrp system is to deliver Avr-like proteins to host cells, identifying these proteins and

determining their function will be key to understanding how *Erwinia chrysanthemi* initiates infection. Recent observations with *Erwinia amylovora* indicate that homologous *avr*-like genes are present in *Erwinia* spp. and *Pseudomonas syringae* (Gaudriault, et al., Mol. Microbiol., 26:1057-1069 (1997); Bogdanove, et al., Proc. Natl. Acad. Sci. USA, 95:1325-1330 (1998), which are hereby incorporated by reference). Specifically, *dspE*, which is required for the pathogenicity of *Erwinia amylovora*, is a homolog of *avrE*, a gene that contributes quantitatively to the virulence of *Pseudomonas syringae* pv *tomato* strain PT23 on tomato and has an Avr phenotype in *Pseudomonas syringae* pv *glycinea* when tested on a variety of soybean cultivars (Lorang, et al., MPMI, 7:508-515 (1994); Lorang, et al., MPMI, 8:49-57 (1995), which are hereby incorporated by reference). The ability of *avrE* to restore the pathogenicity of an *Erwinia amylovora dspE* mutant provides direct evidence that a *Pseudomonas syringae avr* gene can function biologically in an *Erwinia* background. (Bogdanove, et al., Proc. Natl. Acad. Sci. USA, 95:1325-1330 (1998), which is hereby incorporated by reference). Furthermore, DspE-specific antibodies and appropriate *hrp* mutants have been used to establish that *Erwinia amylovora* secretes DspE in a Hrp-dependent manner in culture (Bogdanove, et al., J. Bacteriol. 180:2244-2247 (1998), which is hereby incorporated by reference). however, it is not known whether AvrE can be secreted in culture by *Erwinia amylovora* (or *Pseudomonas syringae*) or whether DspE and AvrE function inside plant cells.

The regulation of the *Erwinia chrysanthemi* Hrp system appears more relaxed in two ways in comparison with host-specific pathogens like *Pseudomonas syringae*. First, the *Erwinia chrysanthemi hrp* genes are not repressed by complex media (which enhances the utility of the system for secretion studies) (Collmer, et al., in "Advances in Molecular Genetics of Plant-Microbe Interactions," Vol. 3 ed. Daniels, M.J. (Kluwer, Dordrecht), pp. 49-56 (1994), which is hereby incorporated by reference). Second, the *Erwinia chrysanthemi* Hrp system does not appear to be gated in culture with respect to the secretion of Avr proteins. Although the *Pseudomonas syringae hrp* cluster carried on pHIR11 enables delivery of AvrB and AvrPto signals (presumably the Avr proteins themselves) to plant cells, it does not direct secretion of these proteins in culture (Figure 5) (Gopalan, et al., Plant Cell, 8:1095-1105 (1996), which is hereby incorporated by reference). Because *Erwinia chrysanthemi* and *Pseudomonas syringae* possess similar Hrp systems (both in group I) (Alfano, et al., J. Bacteriol., 179:5655-5662 (1997);

Kim, et al., "The hrpC and hrpN Operons of *Erwinia chrysanthemi* EC16 are Flanked by plcA and Homologs of Hemolysin/Adhesion Genes and Accompanying Activator/Transporter Genes," MPMI 11(6):563-567 (1998), which are hereby incorporated by reference), comparisons and genetic exchanges between them are likely to be useful for elucidating the mechanisms controlling Avr protein secretion in *Pseudomonas syringae*.

There is yet no direct evidence for the Hrp-mediated transfer of any Avr protein into plant cells, although the indirect evidence for this is particularly compelling with AvrB and AvrPto, as discussed above. The observation that these two proteins can travel the Hrp pathway to the bacterial milieu now provides direct confirmation of the first step in the translocation process. More importantly, the targeting signals controlling secretion and other aspects of the secretion process now can be explored *in vitro*. In this regard, the differing traffic specificities of the type II and type III protein secretion systems of *Erwinia chrysanthemi* are noteworthy, especially since both systems function heterologously in *Escherichia coli*. The cloned cluster of out (type II secretion) genes from *Erwinia chrysanthemi* EC16 directs the secretion of pectate lyase isozymes expressed from *Erwinia chrysanthemi pel* genes but not from *Erwinia carotovora pel* genes (He, et al., Proc. Natl. Acad. Sci. USA, 88:1079-1083 (1991), which is hereby incorporated by reference). This species-specific secretion occurs despite the fact that the Out systems and some of the Pels of these two species are homologous (Lindeberg, et al., Mol. Microbiol., 20:175-190 (1996), which is hereby incorporated by reference). The construction of hybrid Pels has shown that the targeting information controlling species-specific secretion resides in the tertiary structure of these proteins (Lindeberg, et al., J. Bacteriol. 180:1431-1437 (1998), which is hereby incorporated by reference). In contrast, the *Erwinia chrysanthemi* Hrp (type III) system lacks even genus-specificity for its traffic, and the secreted proteins may be devoid of targeting information. This is based on the possibility that targeting information resides in the mRNA encoding the N-termini of these proteins, as has been demonstrated recently for the YopE and YopN proteins secreted by the *Yersinia* type III pathway (Anderson, et al., Science, 278:1140-1143 (1997), which is hereby incorporated by reference). Use of the cloned *Erwinia chrysanthemi* Hrp secretion system should make testing this hypothesis and the identification of targeting signals straightforward.

Escherichia coli heterologously expressing the *Erwinia chrysanthemi* Hrp system also can be used to systematically prospect for genes from *Erwinia chrysanthemi*, *Pseudomonas syringae*, and possibly other bacteria that encode Avr-like effector proteins. However, two factors may limit universal application of this system. First, some Avr-like proteins may require a dedicated chaperone, as has been observed with *Yersinia Yops* (although characteristic gene arrangements and structural properties of the chaperones may help identify them) (Wattiau, et al., Mol. Microbiol., 20:255-262 (1996), which is hereby incorporated by reference). Second, it is not known whether *Escherichia coli*(pCPP2156) will secrete Avr-like proteins derived from pathogens like *Ralstonia solanacearum* and *Xanthomonas* spp., which possess group II Hrp systems (Alfano, et al., J. Bacteriol., 179:5655-5662 (1997), which is hereby incorporated by reference).

Example 5 - Secretion of DspE by the *Erwinia amylovora* Hrp System

The *Erwinia amylovora* DspE (DspA in Gaudriault et al., Mol. Microbiol. 26:1057-1069 (1997), which is hereby incorporated by reference) protein is required for pathogenicity and has homology to AvrE of *Pseudomonas syringae* pv. tomato (Bogdanove et al., Proc. Natl. Acad. Sci. USA 95:1325-1330 (1998); Gaudriault et al., Mol. Microbiol. 26:1057-1069 (1997), which are hereby incorporated by reference). In *Pseudomonas syringae* pv. *glycinia*, DspE acts as an avirulence gene, converting it to avirulence on its host, soybean. In addition, avrE restores pathogenicity to an *Erwinia amylovora* dspE mutant (Bogdanove et al., Proc. Natl. Acad. Sci. USA 95:1325-1330 (1998), which is hereby incorporated by reference). Thus, DspE acts as a virulence protein in host plants and as an avirulence protein in a nonhost plant. To determine whether DspE was secreted by the *Erwinia amylovora* Hrp system, strains Ea273 (the wild-type strain), Ea273-K178 (a hrp secretion mutant), and Ea273dspEΔ1521 (containing a 1,521 bp in-frame deletion in the 3' portion of DspE (Bogdanove et al., J. Bacteriol. 180:2244-2247 (1998); Bogdanove et al., Proc. Natl. Acad. Sci. USA 95:1325-1330 (1998), which are hereby incorporated by reference) were grown in hrp gene-inducing minimal medium. Analysis of the proteins from the cell and supernatant fractions on immunoblots revealed that DspE and DspEΔ1521 were present in the culture supernatant of Ea273 and Ea273dspEΔ1521, respectively (Figure 6). In contrast, no DspE was detected in the supernatant of the hrp mutant (Ea273-K178), but it was

detected within the bacterial cells. Therefore, the DspE protein is secreted in a Hrp-dependent manner.

Example 6 - Construction of a Minimal Hrp Secretion and Regulation System

5 The observations that DspE functions in *Pseudomonas syringae* and that avrE functions in *Erwinia amylovora* raise the possibility that the *Erwinia amylovora* Hrp system functions to secrete other non-*Erwinia* Avr proteins. This was addressed by preparing *Escherichia coli* containing a minimal functional hrp gene cluster from
10 *Erwinia amylovora* strain 321 and two avirulence genes from *Pseudomonas syringae*. The minimal hrp cluster was constructed to avoid interference by the dsp genes linked to the hrp genes in *Erwinia amylovora* (Figure 7). A derivative of pCPP430 with the dsp region deleted, was constructed as follows. The partition region of pCPP9 (the vector portion of pCPP430) and the hrpN gene (from pCPP1084, Wei et al., *Science* 257:85-88
15 (1992), which is hereby incorporated by reference) were cloned into pBluescript II SK- (Stratagene) in the same orientation as in pCPP430. The kanamycin resistance gene from pHP45-Km (Fellay et al., *Gene* 52:147-152 (1987), which is hereby incorporated by reference) was then inserted between the partition region and hrpN. The resulting plasmid was transformed into *Escherichia coli* C2110 (a polA^{ts} strain, ref for Tn3HoHo)
20 containing pCPP430. ColE1 based plasmids such as pBluescript cannot replicate in *Escherichia coli* C2110 at the nonpermissive temperature (42°C), but pSC101 based plasmids such as pCPP9 (the vector portion of pCPP430) can replicate. The resulting strain was grown at 42°C in medium with spectinomycin (for selection of pCPP430) and kanamycin (for selection of the pBluescript construct). The only way that this can occur
25 is if the pBluescript construct integrates into pCPP430 by homologous recombination between either the partition region, or the hrpN gene. When the temperature is lowered to 30°C, a second recombination event takes place. By selecting for bacterial colonies that are resistant to spectinomycin and kanamycin, but sensitive to ampicillin (the selectable marker in pBluescript II SK-) it was possible to find derivatives of pCPP430
30 where the entire region of DNA between the partition region of the vector and the hrpN gene was deleted and replaced by the kanamycin resistance gene. The resulting plasmid was named pCPP431. It contains all of the Hrp secretion and regulatory genes, but none of the putative avirulence genes such as dspE.

To be able to regulate the Hrp secretion system without growing the bacteria in hrp gene-inducing minimal medium a second plasmid was constructed. The hrpL gene is an alternate sigma factor that acts as the global regulator of the rest of the hrp genes. hrpL was cloned into pSU21 (Bartolome et al., Gene 102:75-78 (1991), which is hereby incorporated by reference) under control of the lac promoter to give pCPP1289. When bacteria containing pCPP1289 and pCPP431 are grown in medium containing IPTG the lac promoter is turned on resulting in production of HrpL protein which then turns on the hrp secretion genes that form the Hrp secretion apparatus.

10 **Example 7 - Secretion of *Pseudomonas syringae* Avr Proteins by the *Erwinia amylovora* Hrp System**

Escherichia coli DH5 containing pCPP430 and pCPP1289, or pCPP431 and pCPP1289 was used to test the new systems ability to secrete other proteins.

15 *Escherichia coli* DH5 containing pCPP430 with an insertion in hrcV was used as a secretion-defective control. Derivatives of the avirulence genes avrB and avrPto from *Pseudomonas syringae* were constructed in pFLAG-CTC (Sigma) so that the proteins contained the eight-amino-acid FLAG epitope C-terminal fusions to facilitate detection in immunoblots (Ham et al., "A Cloned *Erwinia Chrysanthemi* Hrp (type III Protein

20 Secretion) System Functions in *Escherichia Coli* to Deliver *Pseudomonas syringae* Avr Signals to Plant Cells and to Secrete Avr Proteins in Culture," Proc. Natl. Acad. Sci. USA, 95(17): 10206-11 (1998), which is hereby incorporated by reference). The *Escherichia coli* strains described above were transformed with the avrB-FLAG and avrPto-FLAG constructs. Each strain was grown in LB medium, induced with IPTG,

25 and the cultures were separated into cell and supernatant fractions. The proteins from the supernatant were concentrated 50 fold by precipitation with 0.1 volumes of 0.15% deoxycholate and 0.1 volumes of 100% TCA. The precipitated proteins were spun down in a centrifuge, rinsed with acetone and dissolved in PAGE gel loading buffer. The proteins were separated by PAGE, blotted, and the avr proteins detected with antiFLAG

30 monoclonal antibodies (Sigma). AvrB-FLAG and AvrPto-FLAG were present in the cell and supernatant fractions from the strains containing the functional hrp gene clusters, pCPP430 or pCPP431 (Figure 8). In contrast, the Avr proteins were present within the cells of the strains containing the hrcV mutation in the hrp cluster, but not present in the supernatant. Thus, the secretion is Hrp-dependent. Assay for the cytoplasmic marker (β -

35 galactosidase revealed nearly imperceptible amounts in the supernatant fractions (not

shown). Thus, cell lysis could not account for the presence of the Avr proteins in the supernatant. Therefore, the *Erwinia amylovora* Hrp system is capable of secreting *Pseudomonas syringae* Avr proteins.

5 **Example 8 - Secretion of Nonbacterial Proteins**

In the Yersinia Yop secretion system the 5'-untranslated portion of the RNA encoding the secreted protein, or the N-terminal portion of the secreted protein contains the signal for secretion through the type secretion system. To test the ability of
10 the *Erwinia amylovora* Hrp secretion system to secrete nonbacterial proteins, a fusion was constructed between the 5'-end of the *Erwinia amylovora* hrpN gene and the mature portion of the human placental alkaline phosphatase gene in the vector pHG165 (Stewart et al., Plasmid 15:172-181 (1986), which is hereby incorporated by reference). The HrpN protein is the main protein secreted by the *Erwinia amylovora* Hrp secretion
15 system. The fusion construct was introduced into *Escherichia coli* DH5 (pCPP431, pCPP1289) and tested for secretion. The construct was not expressed in the *Escherichia coli* strain. The failure to secrete the fusion protein was apparently due to reasons other than the incompatibility of the HrpN secretion signal with the *Erwinia amylovora* Hrp secretion system.

20 Although the invention has been described in detail for the purpose of illustration, it is understood that such detail is solely for that purpose, and variations can be made therein by those skilled in the art without departing from the spirit and scope of the invention which is defined by the following claims.

WHAT IS CLAIMED:

1. A DNA construct comprising:
 - a first DNA molecule encoding a functional type III secretion system;
 - 5 a promoter; and
 - a second DNA molecule encoding a protein or polypeptide capable of being secreted by the type III secretion system, wherein the second DNA molecule is operably coupled to said promoter so that upon introduction of the DNA construct into a host cell, the encoded protein or polypeptide and the type III secretion system are
 - 10 expressed and the encoded protein or polypeptide is secreted.
2. The DNA construct according to claim 1, wherein the encoded protein or polypeptide is a fusion protein comprising:
 - a secretion signal linked by an in-frame gene fusion to a protein or
 - 15 polypeptide of interest.
3. The DNA construct according to claim 2, wherein the fusion protein further comprises:
 - a protease-cleavable amino acid sequence between the secretion signal
 - 20 and the protein or polypeptide of interest.
4. The DNA construct according to claim 2, wherein the secretion signal is a mRNA or a polypeptide fragment of a naturally-occurring protein secreted by the type III secretion system.
- 25 5. The DNA construct according to claim 4, wherein the secretion signal is an N-terminal polypeptide fragment of *Erwinia amylovora* DspE.
6. The DNA construct according to claim 5, wherein the N-terminal
- 30 polypeptide fragment of *Erwinia amylovora* DspE has an amino acid sequence corresponding to SEQ. ID. No. 1.

7. The DNA construct according to claim 6, wherein the N-terminal polypeptide fragment of *Erwinia amylovora* DspE is encoded by a DNA molecule having a nucleotide sequence corresponding to SEQ. ID. No. 2.

5 8. The DNA construct according to claim 1, wherein the first DNA molecule encodes a type III secretion system from *Erwinia amylovora* or *Erwinia chrysanthemi*.

10 9. The DNA construct according to claim 1, wherein the encoded protein or polypeptide is a naturally-secreted protein or polypeptide heterologous to the type III secretion system.

15 10. An expression system comprising an expression vector into which is inserted a heterologous DNA construct of claim 1.

11. A host cell containing a heterologous DNA construct according to claim 1.

20 12. The host cell according to claim 11, wherein the host cell is present in a cell culture.

13. The host cell according to claim 12, wherein the host cell is a prokaryote.

25 14. The host cell according to claim 13, wherein the prokaryote is *Escherichia coli*.

30 15. A system comprising:
a first DNA construct comprising a first DNA molecule encoding a functional type III secretion system and
a second DNA construct comprising promoter operably coupled to a second DNA molecule encoding a protein or polypeptide capable of being secreted by the type III secretion system, wherein upon introduction of the first and second DNA

constructs into a host cell, the encoded protein or polypeptide and the type III secretion system are expressed and the encoded protein or polypeptide is secreted.

16. The system according to claim 15, wherein the first DNA
5 molecule encodes a type III secretion system from *Erwinia amylovora* or *Erwinia chrysanthemi*.

17. The system according to claim 16, wherein the first DNA
10 construct is cosmid pCPP430.

18. The system according to claim 16, wherein the first DNA
construct is cosmid pCPP2156.

19. The system according to claim 15, wherein the encoded protein or
15 polypeptide is a fusion protein comprising:
a secretion signal linked by an in-frame gene fusion to a protein or
polypeptide of interest.

20. The system according to claim 19, wherein the fusion protein
20 further comprises:
a protease-cleavable amino acid sequence between the secretion signal
and the protein or polypeptide of interest.

21. The system according to claim 19, wherein the secretion signal is
25 a mRNA or a polypeptide fragment of a naturally-occurring protein secreted by a type III
secretion system.

22. The system according to claim 21, wherein the secretion signal is
an N-terminal polypeptide fragment of *Erwinia amylovora* DspE.

30 23. The system according to claim 22, wherein the N-terminal
polypeptide fragment of *Erwinia amylovora* DspE has an amino acid sequence
corresponding to SEQ. ID. No. 1.

24. The system according to claim 23, wherein the N-terminal polypeptide fragment of *Erwinia amylovora* DspE is encoded by a DNA molecule having a nucleotide sequence corresponding to SEQ. ID. No. 2.

5

25. The system according to claim 15, wherein the encoded protein or polypeptide is a naturally-secreted protein or polypeptide heterologous to the type III secretion system.

10

26. A host cell containing a heterologous system according to claim 15.

27. The host cell according to claim 26, wherein the host cell is present in a cell culture.

15

28. The host cell according to claim 27, wherein the host cell is a prokaryote.

29. The host cell according to claim 28, wherein the prokaryote is *Escherichia coli*.

20

30. A method of secreting a protein or polypeptide into the environment of a host cell, said method comprising:

introducing into a host cell a DNA construct according to claim 1 under conditions effective to cause expression of the encoded protein or polypeptide and the type III secretion system, wherein the encoded protein or polypeptide is secreted by the host cell into the environment.

25

31. The method according to claim 30, wherein the encoded protein or polypeptide of interest is a fusion protein comprising:

30

a secretion signal linked by an in-frame gene fusion to a protein or polypeptide of interest.

32. The method according to claim 31, wherein the fusion protein further comprises:

a protease-cleavable amino acid sequence between the secretion signal and the protein or polypeptide of interest.

5

33. The method according to claim 31, wherein the secretion signal is a mRNA or a polypeptide fragment of a naturally-occurring protein secreted by a type III secretion system.

10

34. The method according to claim 30, wherein the first DNA molecule encodes a type III secretion system from *Erwinia amylovora* or *Erwinia chrysanthemi*.

15

35. The method according to claim 30, wherein the encoded protein or polypeptide is a naturally-secreted protein or polypeptide heterologous to the type III secretion system.

20

36. The method according to claim 30, wherein the host cell is present in a cell culture.

37. The method according to claim 36, wherein the host cell is a prokaryote.

25

38. The method according to claim 38, wherein the prokaryote is *Escherichia coli*.

39. A method of secreting a protein or polypeptide into the environment of a host cell, said method comprising:

introducing into a host cell a system according to claim 15 under conditions effective to cause expression of the encoded protein or polypeptide and the type III secretion system, wherein the encoded protein or polypeptide is secreted by the host cell into the environment.

30

40. The method according to claim 39, wherein the encoded protein or polypeptide is a fusion protein comprising:

a secretion signal linked by an in-frame gene fusion to a protein or polypeptide of interest.

5

41. The method according to claim 40, wherein the fusion protein further comprises:

a protease-cleavable amino acid sequence between the secretion signal and the protein or polypeptide of interest.

10

42. The method according to claim 40, wherein the secretion signal is a mRNA or a polypeptide fragment of a naturally-occurring protein secreted by a type III secretion system.

15

43. The method according to claim 39, wherein the first DNA molecule encodes a type III secretion system from *Erwinia amylovora* or *Erwinia chrysanthemi*.

20

44. The method according to claim 43, wherein the first DNA construct is cosmid pCPP430.

45. The method according to claim 43, wherein the first DNA construct is cosmid pCPP2156.

25

46. The method according to claim 39, wherein the encoded protein or polypeptide is a naturally-secreted protein or polypeptide heterologous to the type III secretion system.

30

47. The method according to claim 39, wherein the host cell is present in a cell culture.

48. The method according to claim 47, wherein the host cell is a prokaryote.

49. The method according to claim 48, wherein the prokaryote is *Escherichia coli*.

5 50. A method of isolating a protein or polypeptide, said method comprising:

providing a recombinant host cell comprising a first DNA molecule encoding a functional type III secretion system and a second, heterologous DNA molecule having a promoter operably coupled to a nucleic acid sequence encoding a protein or polypeptide capable of being secreted by the type III secretion system;

10 introducing the recombinant host cell into a culture medium, wherein the encoded protein or polypeptide and the type III secretion system are expressed and the encoded protein or polypeptide is secreted into the culture medium; and

isolating the encoded protein or polypeptide from the culture medium.

15 51. The method according to claim 50, wherein the encoded protein or polypeptide is a fusion protein comprising:

a secretion signal linked by an in-frame gene fusion to a protein or polypeptide of interest.

20 52. The method according to claim 51, wherein the fusion protein further comprises:

a protease-cleavable amino acid sequence between the secretion signal and the protein or polypeptide of interest.

25 53. The method according to claim 52, wherein said method further comprises:

introducing into the culture medium, prior to said isolating the encoded protein or polypeptide, a protease effective to cleave the protease-cleavable amino acid sequence.

30 54. The method according to claim 52, wherein said isolating the encoded protein or polypeptide comprises:

isolating the fusion protein from the culture medium;
exposing the isolated fusion protein to a protease effective to cleave the
protease-cleavable amino acid sequence; and
recovering the protein or polypeptide of interest.

5

55. The method according to claim 51, wherein the secretion signal is
a mRNA or a polypeptide fragment of a naturally-occurring protein secreted by a type III
secretion system.

10

56. The method according to claim 50, wherein the type III secretion
system is a type III secretion system from *Erwinia amylovora* or *Erwinia chrysanthemi*.

57. The method according to claim 50, wherein the first DNA
molecule is heterologous to the recombinant host cell.

15

58. The method according to claim 57, wherein the first DNA
molecule is cosmid pCPP430.

59. The method according to claim 57, wherein the first DNA
20 molecule is cosmid pCPP2156.

60. The method according to claim 50, wherein the encoded protein or
polypeptide is a naturally-secreted protein or polypeptide heterologous to the type III
secretion system.

25

61. The method according to claim 50, wherein the host cell is
selected from the group of *Escherichia coli*, *Erwinia amylovora*, and *Erwinia*
chrysanthemi.

30

62. A method of identifying a gene encoding a potential effector
protein or polypeptide comprising:
providing a host cell comprising a DNA molecule encoding a functional
type III secretion system;

introducing into the host cell a candidate gene encoding a protein or polypeptide under conditions effective to express the encoded protein or polypeptide; and

5 determining whether the encoded protein or polypeptide is secreted by the recombinant host cell, wherein secretion of the encoded protein or polypeptide indicates that the gene encodes a potential effector protein or polypeptide.

63. The method according to claim 62, wherein the host cell is selected from the group of *Escherichia coli*, *Erwinia amylovora*, and *Erwinia*
10 *chrysanthemi*.

64. The method according to claim 62, wherein the type III protein secretion system is a type III protein secretion system of *Erwinia amylovora* or *Erwinia*
15 *chrysanthemi*.

65. The method according to claim 62, wherein the DNA molecule is heterologous to the host cell, said method further comprising:
introducing the DNA molecule into the host cell.

20 66. The method according to claim 65, wherein the DNA molecule is cosmid pCPP430 or cosmid pCPP2156.

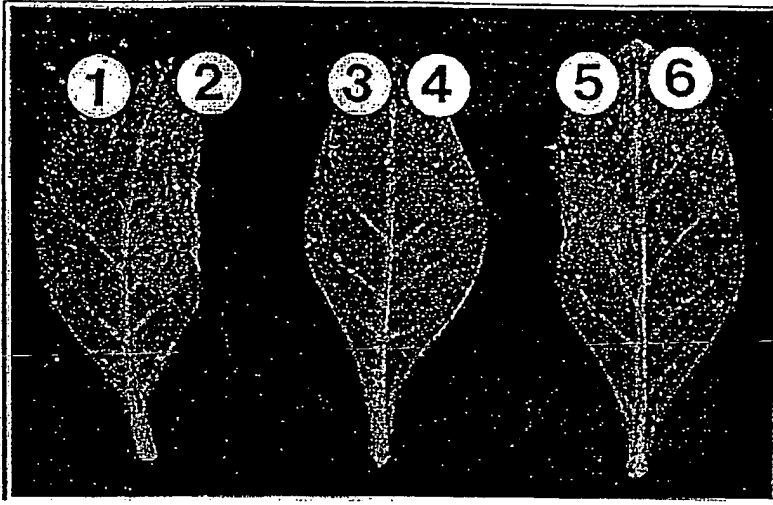
67. The method according to claim 62, wherein the gene is a chimeric gene encoding an epitope tag fused to the encoded protein or polypeptide and said
25 determining comprises:

isolating all protein or polypeptide products secreted by the host cell;
exposing the isolated protein or polypeptide products to an antibody capable of recognizing the epitope tag; and
detecting any antibody bound to the encoded protein or polypeptide.

30 68. The DNA construct according to claim 4, wherein the secretion signal is a mRNA.

69. The system according to claim 21, wherein the secretion signal is a mRNA.

A



B

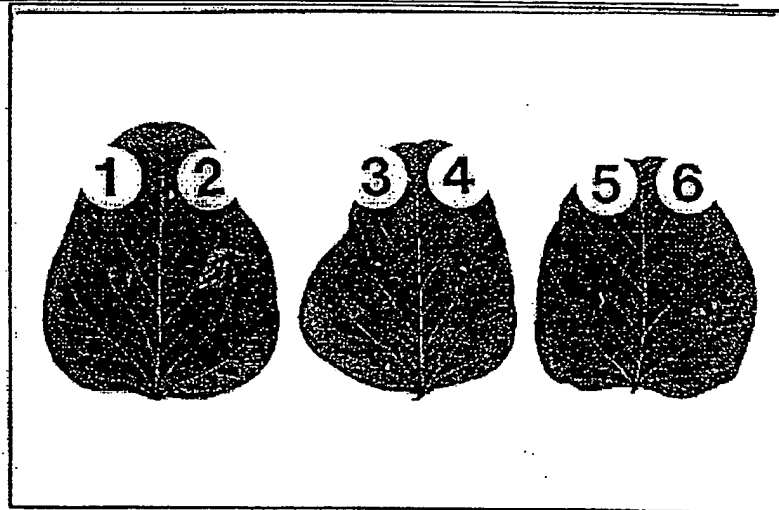


FIGURE 2

FIGURE 3

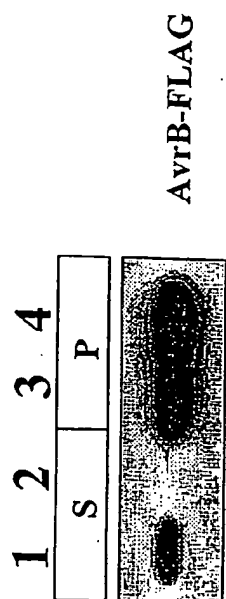


FIGURE 4

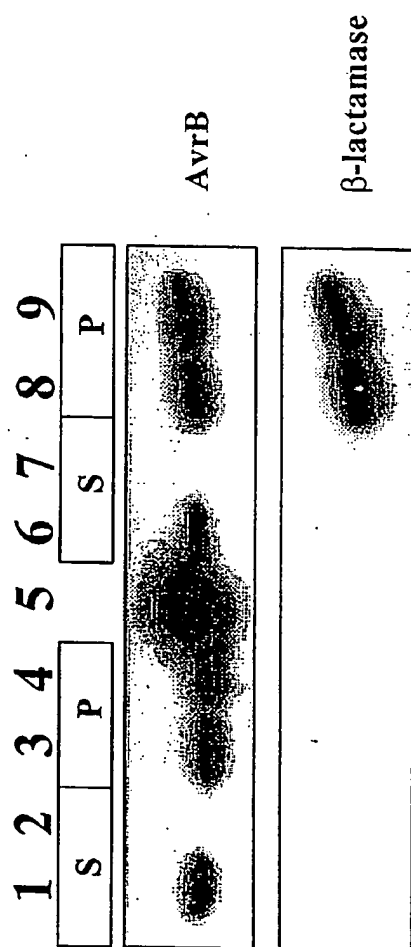
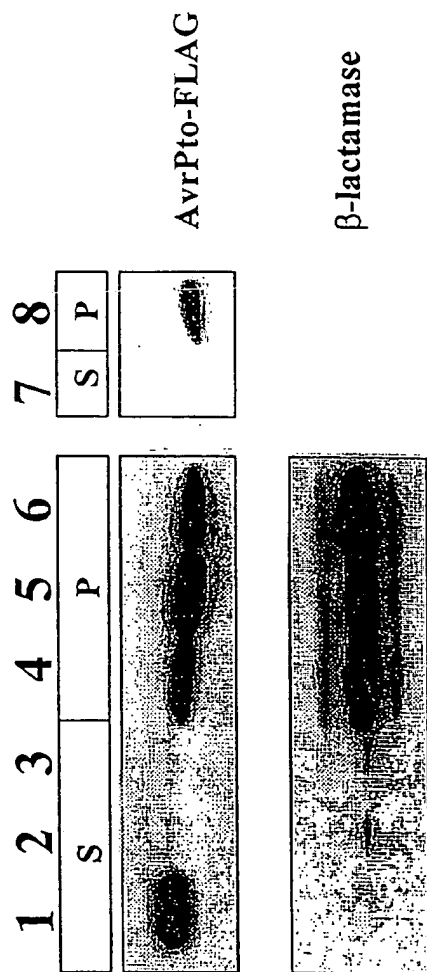
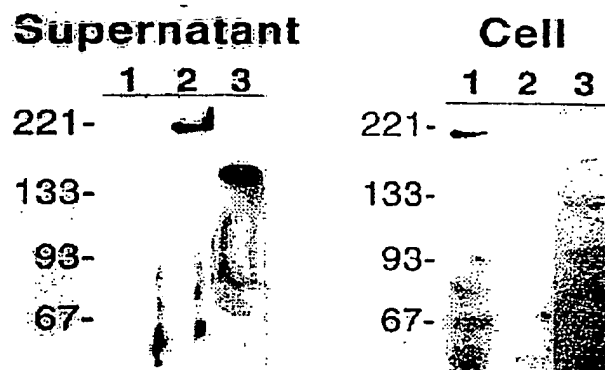


FIGURE 5



**Immunoblots of DspE in Supernatant
and Cell Fractions from *E. amylovora*
Grown in HRP Inducing Medium**

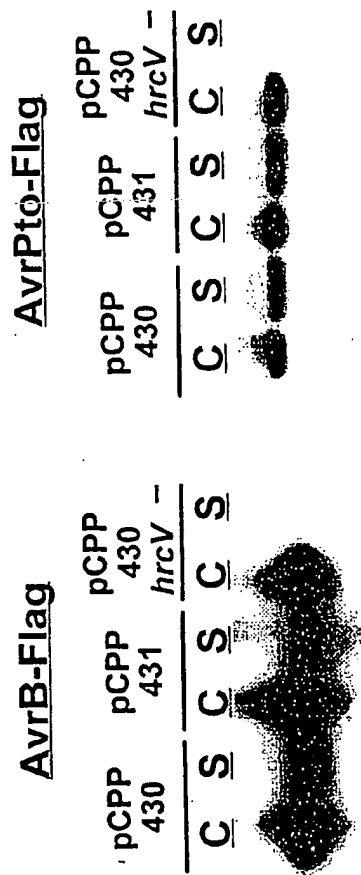


Lane 1. Hrp mutant Ea273-K178.
Lane 2. Wild-type strain Ea273.
Lane 3. Partial *dspE* deletion mutant
strain Ea273*dspE*²¹⁵²¹.

FIGURE 6

FIGURE 8

Immunoblots of Flagged Avr proteins from Cell (C)
and Supernatant (S) Fractions from *E. coli* DH5
with Cosmids containing Various Parts of the *E.*
amylovora Hrp Cluster



SEQUENCE LISTING

<110> Cornell Research Foundation, Inc.

<120> RECOMBINANT CONSTRUCTS AND SYSTEMS FOR SECRETION OF
PROTEINS VIA TYPE III SECRETION SYSTEMS

<130> 19603/2062

<140>

<141>

<150> 60/092,357

<151> 1998-07-10

<160> 6

<170> PatentIn Ver. 2.0

<210> 1

<211> 201

<212> PRT

<213> Erwinia amylovora

<400> 1

Met	Glu	Leu	Lys	Ser	Leu	Gly	Thr	Glu	His	Lys	Ala	Ala	Val	His	Thr
1			5						10					15	

Ala	Ala	His	Asn	Pro	Val	Gly	His	Gly	Val	Ala	Leu	Gln	Gln	Gly	Ser
		20					25						30		

Ser	Ser	Ser	Ser	Pro	Gln	Asn	Ala	Ala	Ala	Ser	Leu	Ala	Ala	Glu	Gly
		35				40					45				

Lys	Asn	Arg	Gly	Lys	Met	Pro	Arg	Ile	His	Gln	Pro	Ser	Thr	Ala	Ala
	50				55					60					

Asp	Gly	Ile	Ser	Ala	Ala	His	Gln	Gln	Lys	Lys	Ser	Phe	Ser	Leu	Arg
65				70					75					80	

Gly	Cys	Leu	Gly	Thr	Lys	Lys	Phe	Ser	Arg	Ser	Ala	Pro	Gln	Gly	Gln
			85						90					95	

Pro	Gly	Thr	Thr	His	Ser	Lys	Gly	Ala	Thr	Leu	Arg	Asp	Leu	Leu	Ala
		100					105						110		

Arg	Asp	Asp	Gly	Glu	Thr	Gln	His	Glu	Ala	Ala	Ala	Pro	Asp	Ala	Ala
	115					120							125		

Arg Leu Thr Arg Ser Gly Gly Val Lys Arg Arg Asn Met Asp Asp Met
 130 135 140

Ala Gly Arg Pro Met Val Lys Gly Gly Ser Gly Glu Asp Lys Val Pro
 145 150 155 160

Thr Gln Gln Lys Arg His Gln Leu Asn Asn Phe Gly Gln Met Arg Gln
 165 170 175

Thr Met Leu Ser Lys Met Ala His Pro Ala Ser Ala Asn Ala Gly Asp
 180 185 190

Arg Leu Gln His Ser Pro Pro His Ile
 195 200

<210> 2

<211> 603

<212> DNA

<213> Erwinia amylovora

<400> 2

atggaattaa aatcactggg aactgaacac aaggcggcag tacacacagc ggcgcacaac 60
 cctgtggggc atggtgttgc cttacagcag ggcagcagca gcagcagccc gcaaaatgcc 120
 gctgcatcat tggcggcaga aggcaaaaat cgtgggaaaa tgccgagaat tcaccagcca 180
 tctactgcgg ctgatggtat cagcgtgtct caccagcaaa agaaatcctt cagtctcagg 240
 ggctgttttg ggacgaaaaa attttccaga tcggcaccgc agggccagcc aggtaccacc 300
 cacagcaaag gggcaacatt gcgcgatctg ctggcgcggg acgacggcga aacgcagcat 360
 gaggcggccg cgccagatgc ggcgcgtttg acccgttcgg gcggcgtcaa acgccgcaat 420
 atggacgaca tggccgggcg gccaatggtg aaagggtggca gcggcgaaga taaggtagca 480
 acgcagcaaa aacggcatca gctgaacaat ttggccaga tgcgccaac gatgttgagc 540
 aaaatggctc acccggtctc agccaacgcc ggcgatcgcc tgcagcattc accgccgcac 600
 atc 603

<210> 3

<211> 19

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: primer

<400> 3

agatctgac aagagacag

19

<210> 4

<211> 19

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: primer

<400> 4

ccgtgtgtat aagagtcag

19

<210> 5

<211> 31

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: primer

<400> 5

gagcgagcat atgggaaata tatgtgtcgg c

31

<210> 6

<211> 31

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: primer

<400> 6

attgtagtcg acttgccagt tacggtacgg g

31

Translation

09/831 907

PATENT COOPERATION TREATY

PCT

INTERNATIONAL PRELIMINARY EXAMINATION REPORT

(PCT Article 36 and Rule 70)

16 26 3
RECEIVED
FEB 14 2002
TECH CENTER 1600/2300

Applicant's or agent's file reference BLOcp598/25P	FOR FURTHER ACTION See Notification of Transmittal of International Preliminary Examination Report (Form PCT/IPEA/416)	
International application No. PCT/FR99/02941	International filing date (day/month/year) 26 November 1999 (26.11.99)	Priority date (day/month/year) 26 November 1998 (26.11.98)
International Patent Classification (IPC) or national classification and IPC C12N 15/16		
Applicant INSTITUT NATIONAL DE LA SANTE ET DE LA RECHERCHE MEDICALE		

1. This international preliminary examination report has been prepared by this International Preliminary Examining Authority and is transmitted to the applicant according to Article 36.
2. This REPORT consists of a total of 6 sheets, including this cover sheet.

☒ This report is also accompanied by ANNEXES, i.e., sheets of the description, claims and/or drawings which have been amended and are the basis for this report and/or sheets containing rectifications made before this Authority (see Rule 70.16 and Section 607 of the Administrative Instructions under the PCT).

These annexes consist of a total of 3 sheets.

3. This report contains indications relating to the following items:

- I ☒ Basis of the report
- II ☐ Priority
- III ☐ Non-establishment of opinion with regard to novelty, inventive step and industrial applicability
- IV ☐ Lack of unity of invention
- V ☒ Reasoned statement under Article 35(2) with regard to novelty, inventive step or industrial applicability; citations and explanations supporting such statement
- VI ☒ Certain documents cited
- VII ☐ Certain defects in the international application
- VIII ☒ Certain observations on the international application

Date of submission of the demand 16 June 2000 (16.06.00)	Date of completion of this report 06 February 2001 (06.02.2001)
Name and mailing address of the IPEA/EP	Authorized officer
Facsimile No.	Telephone No.

INTERNATIONAL PRELIMINARY EXAMINATION REPORT

International application No.

PCT/FR99/02941

I. Basis of the report

1. With regard to the elements of the international application:*

- ☐ the international application as originally filed
- ☒ the description:
 pages 1-12, as originally filed
 pages _____, filed with the demand
 pages 13, filed with the letter of 23 January 2001 (23.01.2001)
- ☒ the claims:
 pages _____, as originally filed
 pages _____, as amended (together with any statement under Article 19
 pages _____, filed with the demand
 pages 1-14, filed with the letter of 23 January 2001 (23.01.2001)
- ☒ the drawings:
 pages 1/8-8/8, as originally filed
 pages _____, filed with the demand
 pages _____, filed with the letter of _____
- ☒ the sequence listing part of the description:
 pages 1-10, as originally filed
 pages _____, filed with the demand
 pages _____, filed with the letter of _____

2. With regard to the language, all the elements marked above were available or furnished to this Authority in the language in which the international application was filed, unless otherwise indicated under this item.

These elements were available or furnished to this Authority in the following language _____ which is:

- ☐ the language of a translation furnished for the purposes of international search (under Rule 23.1(b)).
- ☐ the language of publication of the international application (under Rule 48.3(b)).
- ☐ the language of the translation furnished for the purposes of international preliminary examination (under Rule 55.2 and/or 55.3).

3. With regard to any nucleotide and/or amino acid sequence disclosed in the international application, the international preliminary examination was carried out on the basis of the sequence listing:

- ☒ contained in the international application in written form.
- ☐ filed together with the international application in computer readable form.
- ☐ furnished subsequently to this Authority in written form.
- ☐ furnished subsequently to this Authority in computer readable form.
- ☐ The statement that the subsequently furnished written sequence listing does not go beyond the disclosure in the international application as filed has been furnished.
- ☐ The statement that the information recorded in computer readable form is identical to the written sequence listing has been furnished.

4. ☐ The amendments have resulted in the cancellation of:

- ☐ the description, pages _____
- ☐ the claims, Nos. _____
- ☐ the drawings, sheets/fig _____

5. ☐ This report has been established as if (some of) the amendments had not been made, since they have been considered to go beyond the disclosure as filed, as indicated in the Supplemental Box (Rule 70.2(c)).**

* Replacement sheets which have been furnished to the receiving Office in response to an invitation under Article 14 are referred to in this report as "originally filed" and are not annexed to this report since they do not contain amendments (Rule 70.16 and 70.17).

** Any replacement sheet containing such amendments must be referred to under item 1 and annexed to this report.

INTERNATIONAL PRELIMINARY EXAMINATION REPORT

International application No.
PCT/FR99/02941

VI. Certain documents cited

1. Certain published documents (Rule 70.10)

Application No. Patent No.	Publication date (day/month/year)	Filing date (day/month/year)	Priority date (valid claim) (day/month/year)
WO 99/35266			
WO 00/00610			

2. Non-written disclosures (Rule 70.9)

Kind of non-written disclosure	Date of non-written disclosure (day/month/year)	Date of written disclosure referring to non-written disclosure (day/month/year)

INTERNATIONAL PRELIMINARY EXAMINATION REPORT

International application No.
PCT/FR 99/02941

V. Reasoned statement under Article 35(2) with regard to novelty, inventive step or industrial applicability; citations and explanations supporting such statement

1. Statement

Novelty (N)	Claims	1, 2, 4, 7-14	YES
	Claims	3, 5, 6	NO
Inventive step (IS)	Claims	1, 2, 4, 7-14	YES
	Claims	3, 5, 6	NO
Industrial applicability (IA)	Claims	1-14	YES
	Claims		NO

2. Citations and explanations

1. Reference is made to the following document:

D1: Database EMBL ID AF172174
Accession Number Z98884

2. D1 describes the sequence of a genomic DNA fragment. This sequence comprises a coding sequence disclosing 99% identical to the cDNA sequence of SEQ ID NO:4. The sequence of SEQ ID NO:4 comprises the sequences coding for polypeptides SEQ ID NO:1-3. Consequently, in view of Document D1, the subject matter of Claim 3 is not novel. In addition, it should be noted that, by virtue of the wording of Claim 3, independently of Document 1, this claim could not be acknowledged as novel. Indeed, Claim 3 refers to oligonucleotides from a specific sequence. Claim 3 does not define, in particular, the length of the claimed oligonucleotides. Hence, in the absence of such an indication, these oligonucleotides can, in particular, consist of only some nucleotides. Such oligonucleotides are too short to be specific to a given sequence. Consequently, the fragments of Claim 3 include oligonucleotides that are not specific to a given

sequence, some being known in the prior art. Claim 3 specifies that the oligonucleotides constitute probes or primers. This does not correspond to any technical feature enabling these oligonucleotides to be characterised, in particular, their length or their specificity for a particular sequence, and thus cannot restore their novelty. For instance, the hexanucleotides are not specific for a given sequence, are known in the prior art, and can act as a probe or primer. Vectors containing these known oligonucleotides, as well as cells transformed by such vectors, are known in the prior art. Claims 3, 5, and 6, therefore, do not fulfil the requirements of PCT Article 33(2).

VIII. Certain observations on the international application

The following observations on the clarity of the claims, description, and drawings or on the question whether the claims are fully supported by the description, are made:

1. The reference to subjects introduced by the expression "in particular" (Claim 11) is given by way of indicative example and is in no way exhaustive, and cannot be considered to limit the subject matter of the claim in question.
2. Claim 6 lacks clarity by virtue of the wording "cells transformed by at least one nucleic acid fragment". Indeed, the cells are usually transformed by a vector comprising one or more specific fragments, and not by nucleic acid fragments (PCT Article 6).